

92847372

EP 89 / 2055



REC'D 18 MAY 1999

WIPO PCT

Bescheinigung

Herr Professor Dr. André Rosenthal Institut für Molekulare Biotechnologie in Jena/
Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Nucleinsäuremoleküle, codierend Proteine, die die Knochen-
entwicklung beeinflussen"

am 27. März 1998 beim Deutschen Patent- und Markenamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ur-
sprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patent- und Markenamt vorläufig die Symbole
C 07 K, C 12 N und C 07 H der Internationalen Patentklassifikation erhalten.

München, den 26. April 1999

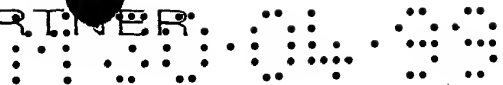
Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

Faust

Aktenzeichen: 198 13 799.0



PATENTANWÄLTE
EUROPEAN PATENT ATTORNEYS
EUROPEAN TRADEMARK ATTORNEYS
Dr. VOLKER VOSSIUS, Dipl.-Chem.

(bis 1992; danach in anderer Kanzlei)
Dr. PAUL TAUCHNER, Dipl.-Chem.
Dr. DIETER HEUNEMANN, Dipl.-Phys.
Dr. PETER A. RAUH, Dipl.-Chem.
Dr. GERHARD HERMANN, Dipl.-Phys.
JOSEF SCHMIDT, Dipl.-Ing.
Dr. HANS-RAINER JAENICHEN, Dipl.-Biol.
Dr. ALEXA VON UEXKÜLL, M. Sc.
Dr. RUDOLF WEINBERGER, Dipl.-Chem.
Dr. WOLFGANG BUBLAK, Dipl.-Chem.
AXEL STELLBRINK, Dipl.-Ing.
Dr. JOACHIM WACHENFELD, (Biol.)

EUROPEAN PATENT ATTORNEY
Dr. RENATE BARTH, Dipl.-Chem.

RECHTSANWÄLTE
HELGA TREMMEL
BARBARA GUGGENMOS, Dipl.-Chem.

SIEBERTSTRASSE 4
81675 MÜNCHEN

POB 860767
81634 MÜNCHEN
GERMANY

TELEFON: +49-89-4 13 04-0
FAX G3: +49-89-4 13 04-111
FAX G4: +49-89-4 13 04-101

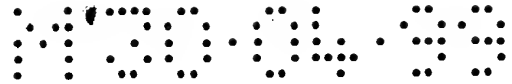
u.Z.: C 1465 DE

Prof. Dr. André Rosenthal
Institut für Molekulare Biotechnologie
Abteilung für Genomanalyse
D-07745 Jena, DE

Nucleinsäuremoleküle, codierend Proteine, die die Knochenentwicklung beeinflussen

Die vorliegende Erfindung betrifft Nucleinsäuremoleküle codierend Proteine, die die Knochenentwicklung in Säugern beeinflussen, die codierten Proteine, sowie diagnostische und pharmazeutische Zusammensetzungen enthaltend derartige Nucleinsäuremoleküle oder Proteine. Die Erfindung betrifft ferner transgene nicht-menschliche Säuger, die mit den beschriebenen Nucleinsäuremolekülen transformiert sind oder die eine veränderte Expression der beschriebenen Proteine aufweisen.

Beim Menschen sind eine Reihe von Erbkrankheiten bekannt, die zu Wachstums- und Entwicklungsstörungen der Knochen führen. Hierzu zählen beispielsweise spondyloepiphyseäre Dysplasien und Achondroplasie. Die genauen genetischen Ursachen derartiger Störungen sind in der Regel nicht geklärt und

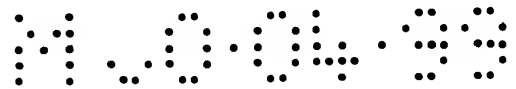


therapeutische Ansätze oder diagnostische Verfahren zur Früherkennung stehen größtenteils nicht zur Verfügung. Die Aufklärung der Ursachen solcher Wachstums- und Entwicklungsstörungen sowie die Bereitstellung möglicher therapeutischer Ansätze und Diagnostikverfahren zur frühzeitigen Erkennung solcher Störungen erfordert die Identifizierung und Isolierung von Genen, die an der Regulation entsprechender Wachstums- und Entwicklungsprozesse beteiligt sind.

Der vorliegenden Erfindung liegt somit die Aufgabe zugrunde, Nucleinsäuremoleküle zur Verfügung zu stellen, deren Expressionsprodukt Wachstums- und Entwicklungsprozesse, insbesondere im Zusammenhang mit Knochen, bei Tieren und Menschen beeinflusst.

Diese Aufgabe wird durch die Bereitstellung der in den Patentansprüchen bezeichneten Ausführungsformen gelöst.

Somit betrifft die vorliegende Erfindung Nucleinsäuremoleküle, die eine Nucleotidsequenz umfassen, die die in Seq ID No. 2 oder in Seq ID No. 4 dargestellte Aminosäuresequenz codiert, sowie Nucleinsäuremoleküle, die die in Seq ID No. 1 oder Seq ID No. 3 dargestellte Nucleotidsequenz umfassen. Es wurde überraschenderweise gefunden, daß das von einem solchen Nucleinsäuremolekül codierte Protein ein Protein ist, dessen Inaktivierung in Säugern zu einer Verlängerung der Knochen mit Ausnahme der Schädelknochen führt. Aufgefunden wurden derartige Nucleinsäuremoleküle im Zusammenhang mit der Erzeugung einer sogenannten transgenen "Donor"-Maus, d.h. einer Maus, die als Spender für ein künstliches Protein dienen sollte. Dieses künstliche Protein sollte in bestimmten Geweben der "Donor"-Maus exprimiert werden, ohne jedoch in dieser Maus eine Funktion zu haben. Erst nach Kreuzung der Donor-Maus mit einer geeigneten transgenen Empfänger-Maus sollte das Protein wirksam werden und bestimmte Gene der Empfänger-Maus aktivieren. Transgene Donor-Mäuse sind schon des öfteren erzeugt worden. Normalerweise zeigen sie keinen



Phänotyp, da das künstliche Gen einfach in befruchtete Eizellen gespritzt wird und sich rein zufällig in irgendeinem Bereich des murinen Genoms integriert. Da lediglich etwa 5 % des Genoms codierend sind, ist die Wahrscheinlichkeit, daß ein Defekt in einem essentiellen Gen verursacht wird entsprechend gering. Hinzu kommt, daß das Säugergenom diploid ist, d.h. alle Gene sind in doppelter Ausführung vorhanden. Daher sind die meisten Mutationen rezessiv, d.h. sie kommen nicht zur Ausprägung: dem mutierten Gen steht eine voll funktionsfähige Kopie gegenüber, die den erzeugten Defekt kompensieren kann.

Überraschenderweise zeigte die erzeugte Donormaus einen extrem auffälligen Phänotyp: alle Knochen (mit Ausnahme der des Schädels) sind um das 1,3- bis 1,5-fache verlängert. Infolgedessen ist die transgene Maus etwa 1,5 mal länger als der entsprechende Wildtyp (siehe Fig. 1). Dieser Phänotyp ist dominant und wird stabil weitervererbt, d.h. bei Kreuzung einer transgenen Mutane mit einer gesunden Wildtyp-Maus zeigen 50 % der Nachkommen den oben beschriebenen Phänotyp.

Die genetische Analyse dieser Maus ergab, daß durch die Insertion der DNA für das in der Maus zu produzierende künstliche Protein in das Genom der Maus ein Gen inaktiviert wurde. Um herauszufinden, welches Gen (oder welche Gene) für den beobachteten Phänotyp verantwortlich sind, wurde der mutierte Bereich des Genoms der transgenen Maus in Bakterien subcloniert. Die Lokalisierung des mutierten Bereiches im Genom der Maus und die anschließende Subclonierung waren möglich, weil die Nucleotidsequenz des eingeführten künstlichen Gens bekannt war und diese Information in entsprechenden molekularbiologischen Experimenten genutzt werden konnte.

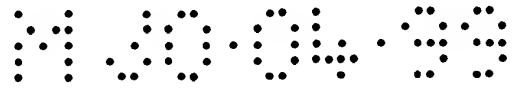
Zur Identifizierung des Gens, das im folgenden LOBO-Gen ("long bones") genannt wird, wurden 6 kb aus dem subclonierten Bereich der transgenen Maus sowie 87 kb aus der entsprechenden homologen Region der Wildtyp-Maus sequenziert. Eine detaillierte Computeranalyse der Sequenzdaten führte zur Identifizierung eines Gens, das aus mindestens 8 codierenden



Abschnitten ("Exons") besteht und mindestens 60 000 Basen lang ist, wahrscheinlich aber sehr viel länger. Der bisher identifizierte, codierende Bereich trägt die Information für 393 Aminosäuren (siehe Seq ID No. 2). Aufgrund der erhaltenen murinen Sequenzdaten wurde eine DNA-Sonde konstruiert, mit deren Hilfe ein humaner P1 Clon isoliert wurde, der das menschliche LOBO-homologe Gen trägt. Die Sequenz der bisher isolierten und identifizierten codierenden Bereiche (Exons) dieses Gens ist in Seq ID No. 3 dargestellt, ebenso wie die davon abgeleitete Aminosäuresequenz. Ein Sequenzvergleich der erhaltenen murinen und menschlichen Sequenzen mit bekannten Sequenzen ergab, daß das codierte Protein eine gewisse Homologie zu einem Protein aus *C. elegans* hat. Da das Protein aus *C. elegans* am N-Terminus ca. 400 Aminosäuren länger ist als das von den bisher in Maus identifizierten Sequenzen codierte Protein, ist zu vermuten, daß das murine LOBO-Gen weitere Sequenzen umfaßt und bisher nur partiell sequenziert wurde. Mit Hilfe der durch die vorliegende Erfindung zur Verfügung gestellten Sequenzdaten ist es jedoch möglich mittels Standardverfahren, z.B. chromosomal walking, die noch fehlenden Bereiche des Gens zu isolieren und zu charakterisieren.

Aus dem oben gesagten ergibt sich, daß die erfindungsgemäßen Nucleinsäuremoleküle ein Protein codieren, dessen Inaktivierung in Tieren, vorzugsweise in Vertebraten, bevorzugt in Säugern und besonders bevorzugt in Maus zu einer Verlängerung der Knochen mit Ausnahme der Schädelknochen führen. Eine Verlängerung bedeutet dabei vorzugsweise eine Verlängerung um einen Faktor von mindestens 1,2, bevorzugt um einen Faktor von 1,3 und besonders bevorzugt um einen Faktor im Bereich von 1,3 bis 1,5.

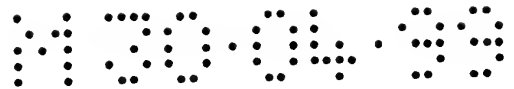
Der Begriff "Inaktivierung" bedeutet dabei, daß die Expression des Proteins im Vergleich zum Wildtyp verringert ist, vorzugsweise um mindestens 50 % und besonders bevorzugt total reprimiert ist. Die Analyse der Mutation im Genom der oben beschriebenen Donormaus ergab, daß sich die Insertion



des künstlichen Gens innerhalb eines Introns des LOBO-Gens befindet und zur Deletion von 10 Basenpaaren geführt hat. Letzteres sollte im Intron kein Problem bereiten, da dieser Bereich ohnehin nicht codiert. Es ist daher anzunehmen, daß es aufgrund der künstlichen DNA-Insertion zur Störung bei der Reifung der mRNA kommt ("splicing"), da das künstlich eingeführte Gen Splicing-Signale enthält. Dies führt vermutlich zu einem sogenannten "aberranten splicing". Infolgedessen wird die Bildung einer funktionsfähigen mRNA verhindert und das entsprechende Protein kann nicht produziert werden. Tatsächlich hat die experimentelle Überprüfung der LOBO-Expression (durch "Northern Blot") ergeben, daß in heterozygoten LOBO-Mäusen nur noch etwa die Hälfte an mRNA produziert wird, im Vergleich zur Wildtyp-Maus. In homozygoten LOBO-Mäusen kann überhaupt keine LOBO-mRNA mehr nachgewiesen werden. Es ist daher anzunehmen, daß die Mutation in der transgenen LOBO-Maus die Genexpression auf post-transkriptionaler Ebene abschaltet. Offenbar sinkt dann bereits in den heterozygoten Mäusen die Menge an produziertem LOBO-Protein unter einen kritischen Schwellenwert, was dann zu dem beobachteten, dominanten Phänotyp führt.

Der Begriff "Inaktivierung" bedeutet daher im Rahmen der vorliegenden Erfindung vorzugsweise, daß die Menge an Transkripten, die das beschriebene Protein codieren, in den Zellen im Vergleich zu Zellen von entsprechenden Wildtyp-Tieren um mindestens 50 % verringert ist, vorzugsweise um mindestens 70 %, besonders bevorzugt um mindestens 90 %. In einer ganz besonders bevorzugten Ausführungsform bedeutet "Inaktivierung", daß keinerlei Transkripte mehr nachgewiesen werden können, die das beschriebene Protein codieren. Die Menge an Transkripten kann nach dem Fachmann bekannten Techniken nachgewiesen werden, beispielsweise durch Northern Blot-Analyse.

Vorzugsweise führt die Inaktivierung des durch die erfindungsgemäßen Nucleinsäuremoleküle codierten Proteins in Mäu-



sen weiterhin zu mindestens einer der folgenden Veränderungen:

- (a) die Wandung der Knochen ist mindestens doppelt so dick wie bei Wildtyp-Mäusen (siehe Figur 4);
- (b) das Knochenmark enthält wesentlich weniger stabilisierende Elemente als Knochen von Wildtyp-Mäusen (siehe Figur 4);
- (c) die Knochen sind brüchiger als Knochen von Wildtyp-Mäusen;
- (d) die Lebenserwartung ist drastisch verkürzt, beträgt maximal 40 Wochen und durchschnittlich ca. 25 Wochen (bei Wildtyp-Mäusen liegt die durchschnittliche Lebenserwartung bei 1 bis 2 Jahren).

Der Vergleich der Aminosäuresequenz des durch die erfindungsgemäßen Nucleinsäuremoleküle codierten Proteins zeigte, daß die Aminosäuresequenz Bereiche aufweist, die zwischen Mensch und Maus stark konserviert sind und die auch in Proteinen aus *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, aus einem Bakterium der Gattung *Leuconostoc* und aus *Caenorhabditis elegans* vorkommen. Von dem verwandten Protein aus *S. cerevisiae*, Dis3, ist bekannt, daß es eine Rolle in der Zellzykluskontrolle spielt. Die Funktion der anderen homologen Proteine ist bisher nicht bekannt. Das Vorhandensein von verwandten Proteinen in Bakterien als auch in Hefen und *C. elegans* läßt vermuten, daß die erfindungsgemäßen Nucleinsäuremoleküle ein Protein codieren, daß eine zentrale Rolle auf zytologischer Ebene spielt. Das in der vorliegenden Erfindung beschriebene Nucleinsäuremolekül aus der Maus scheint eine wesentliche Rolle in der murinen Individualentwicklung zu spielen, denn die Verlängerung der Knochen in der Mutante ist bereits direkt nach der Geburt sichtbar und wird dann mit zunehmendem Alter der Tiere immer dramatischer. Das übersteigerte Knochenwachstum könnte endokrinologische Ursachen haben, d.h. das LOBO-Gen könnte ein Wachstumshormon produzieren oder die Produktion eines solchen Hormons durch Steuerung anderer Gene kontrollieren oder

selbst ein Rezeptor für Wachstumshormone sein. All dies ist aber unwahrscheinlich, da bisher keinerlei Sequenzhomologien zu einer der genannten Proteinklassen gefunden wurde. Dagegen scheint es aufgrund der oben gezeigten Homologie zum Dis3-Gen aus *S. cerevisiae* wahrscheinlicher, daß durch die Mutation im LOBO-Gen die Steuerung der Zellteilung (Mitose) gestört wird und es daher zu verändertem Wachstum kommt.

Somit ist die Bereitstellung der erfindungsgemäßen Nucleinsäuremoleküle sowohl wissenschaftlich als auch klinisch von großer Bedeutung. Seine weitere Erforschung könnte zum einen helfen, die Zellzykluskontrolle noch besser zu verstehen. Dies ist insbesondere für die Krebsforschung wichtig. Zum anderen könnten die erfindungsgemäßen Nucleinsäuremoleküle für humane Wachstumsstörungen verantwortlich sein, die nicht ernährungs- oder hormonbedingt sind.

Die vorliegende Erfindung betrifft auch Nucleinsäuremoleküle, deren komplementärer Strang mit einem der obenbeschriebenen erfindungsgemäßen Nucleinsäuremoleküle hybridisiert und die ein Protein mit den obengenannten Eigenschaften codieren.

Der Begriff "Hybridisierung" bedeutet im Rahmen der vorliegenden Erfindung eine Hybridisierung unter konventionellen Hybridisierungsbedingungen, vorzugsweise unter stringenten Bedingungen, wie sie beispielsweise in Sambrock et al., *Molecular Cloning, A Laboratory Manual*, 2. Aufl. (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) beschrieben sind. Dabei bedeutet "stringente Bedingungen", daß eine Hybridisierung nur erfolgt, wenn eine Sequenzidentität von mindestens 90 %, vorzugsweise von mindestens 95 % und besonders bevorzugt von mindestens 97 % über die gesamte Länge vorliegt. Nucleinsäuremoleküle, die mit den erfindungsgemäßen Nucleinsäuremolekülen hybridisieren, können prinzipiell aus jedem beliebigen tierischen Organismus stammen, der ein derartiges Protein exprimiert. Vorzugsweise sind es Moleküle, die entsprechende Proteine aus höheren tierischen Organismen codieren, bevorzugt aus Vertebraten,

besonders bevorzugt aus Säugern und insbesondere aus Maus oder Mensch.

Nucleinsäuremoleküle, die mit den erfindungsgemäßen Molekülen hybridisieren, können z.B. aus genomischen oder aus cDNA-Bibliotheken isoliert werden. Die Identifizierung und Isolierung derartiger Nucleinsäuremoleküle kann dabei unter Verwendung der erfindungsgemäßen Nucleinsäuremoleküle oder Teile dieser Moleküle bzw. der reversen Komplemente dieser Moleküle erfolgen, z.B. mittels Hybridisierung nach Standardverfahren (siehe z.B. Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2. Aufl. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) oder durch Amplifikation mittels PCR.

Als Hybridisierungsprobe können z.B. Nucleinsäuremoleküle verwendet werden, die exakt die oder im wesentlichen die unter Seq ID No. 1 oder 3 angegebene Nucleotidsequenz oder Teile dieser Sequenz aufweisen. Bei den als Hybridisierungsprobe verwendeten Fragmenten kann es sich auch um synthetische Fragmente handeln, die mit Hilfe der gängigen Synthesetechniken hergestellt wurden und deren Sequenz im wesentlichen mit der eines erfindungsgemäßen Nucleinsäuremoleküls übereinstimmt. Hat man Gene identifiziert und isoliert, die mit den erfindungsgemäßen Nucleinsäuresequenzen hybridisieren, sollte eine Bestimmung der Sequenz und eine Analyse der Eigenschaften der von dieser Sequenz codierten Proteine erfolgen.

Die mit den erfindungsgemäßen Nucleinsäuremolekülen hybridisierenden Moleküle umfassen insbesondere Fragmente, Derivate und allelische Varianten der oben beschriebenen Nucleinsäuremoleküle, die ein Protein mit den vorstehend beschriebenen Eigenschaften codieren. Der Ausdruck Derivat bedeutet in diesem Zusammenhang, daß die Sequenzen dieser Moleküle sich von den Sequenzen der oben beschriebenen Nucleinsäuremoleküle an einer oder mehreren Positionen unterscheiden und einen hohen Grad an Homologie zu diesen Sequenzen aufweisen. Homologie bedeutet dabei eine Sequenzidentität auf Aminosäureebene über die gesamte Länge von mindestens 70 %, insbe-



sondere eine Identität von mindestens 80 %, vorzugsweise über 90 %, besonders bevorzugt über 95 % und insbesondere von mindestens 97 %. Vorzugsweise bedeutet Homologie ferner eine Sequenzidentität auf der Nucleinsäuresequenzebene von mindestens 60 %, vorzugsweise mindestens 70 %, besonders bevorzugt mindestens 85 % und insbesondere bevorzugt von mindestens 95 %. Die Abweichungen zu den oben beschriebenen Nucleinsäuremolekülen können dabei z.B. durch Deletion, Addition, Substitution, Insertion oder Rekombination entstanden sein.

Homologie bedeutet ferner, daß funktionelle und/oder strukturelle Äquivalenz zwischen den betreffenden Nucleinsäuremolekülen oder den durch sie codierten Proteinen, besteht. Bei den Nucleinsäuremolekülen, die homolog zu den oben beschriebenen Molekülen sind und Derivate dieser Moleküle darstellen, handelt es sich in der Regel um Variationen dieser Moleküle, die Modifikationen darstellen, die dieselbe biologische Funktion ausüben. Es kann sich dabei sowohl um natürlicherweise auftretende Variationen handeln, beispielsweise um Sequenzen aus anderen Tierarten, oder um Mutationen, wobei diese Mutationen auf natürliche Weise aufgetreten sein können oder durch gezielte Mutagenese eingeführt wurden. Ferner kann es sich bei den Variationen um synthetisch hergestellte Sequenzen handeln. Bei den allelischen Varianten kann es sich sowohl um natürlich auftretende Varianten handeln, als auch um synthetisch hergestellte oder durch rekombinante DNA-Techniken erzeugte Varianten.

Die von den verschiedenen Varianten der erfindungsgemäßen Nucleinsäuremoleküle codierten Proteine weisen bestimmte gemeinsame Charakteristika auf. Dazu können z.B. biologische Aktivität, Molekulargewicht, immunologische Reaktivität, Konformation etc. gehören, sowie physikalische Eigenschaften wie z.B. das Laufverhalten in Gelelektrophoresen, chromatographisches Verhalten, Sedimentationskoeffizienten, Löslichkeit, spektroskopische Eigenschaften, Stabilität; pH-Optimum, Temperatur-Optimum etc.



Die Proteine, die von den erfindungsgemäßen Nucleinsäuremolekülen codiert werden, weisen vorzugsweise dieselbe biologische Funktion oder Aktivität auf wie oben für das murine Protein beschrieben, d.h. bei Inaktivierung dieser Proteine in Tieren kommt es zur oben beschriebenen Veränderung der Knochenentwicklung.

Besonders bevorzugt weist das durch ein erfindungsgemäßes Nucleinsäuremolekül codierte Protein mindestens eine der beiden folgenden Konsensussequenzen auf.

Konsensus 1:

EFMLLANXXVAXXIXXXFPXXALLRRHXXP

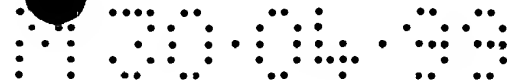
Konsensus 2:

HZALNVXXZTHFTSPIRRZXDVIVHRLLAALGY

Die vorliegende Erfindung betrifft ferner Nucleinsäuremoleküle deren Sequenz von der Sequenz eines oben beschriebenen Nucleinsäuremoleküls aufgrund der Degeneration des genetischen Codes abweicht.

Die erfindungsgemäßen Nucleinsäuremoleküle können beliebige Nucleinsäuremoleküle sein, insbesondere DNA- oder RNA-Moleküle, beispielsweise cDNA, genomische DNA, mRNA etc. Sie können natürlich vorkommende Moleküle sein, oder durch gentechnische oder chemische Syntheseverfahren hergestellte Moleküle.

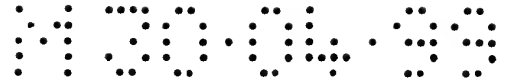
Beispiele für genomische Sequenzen der Maus und des Menschen sind in Seq ID No. 5, 6 und 7 dargestellt. Mit Hilfe von "fluorescent in situ hybridization" (Fish) an kompletten murinen Metaphasechromosomen wurde das murine Gen in der Bande 1D auf dem Chromosomen 1 der Maus lokalisiert. Diese Bande ist mit der Bande 2q35 auf dem humanen Chromosom 2 homolog. In diesem Abschnitt befindet sich auch ein Gen für alkalische Phosphatase, dessen Position in der Literatur genau bekannt ist. Die Analyse der genomischen Sequenzen aus Maus und Mensch, die ein erfindungsgemäßes Nucleinsäuremolekül tragen, ergab, daß sich in beiden Fällen das Gen für die alkalische Phosphatase ca. 20 kb stromabwärts des LOBO-Gens



befindet, so daß dessen chromosomale Lokalisation sehr genau angegeben werden kann. Mit Hilfe der in der vorliegenden Erfindung offenbarten Nucleinsäuremoleküle ist es dem Fachmann möglich, mittels bekannter Verfahren homologe Sequenzen aus anderen Organismen, insbesondere Säugern zu isolieren.

Weiterhin betrifft die Erfindung Vektoren, insbesondere Plasmide, Cosmide, Viren, Bacteriophagen und andere in der Gentechnik gängige Vektoren, die die oben beschriebenen erfindungsgemäßen Nucleinsäuremoleküle enthalten.

In einer bevorzugten Ausführungsform sind die in den Vektoren enthaltenen Nucleinsäuremoleküle verknüpft mit regulatorischen Elementen, die die Expression in prokaryontischen oder eukaryontischen Zellen gewährleisten. Der Begriff "Expression" kann dabei Transkription als auch Transkription und Translation bedeuten. Regulatorische Elemente umfassen dabei insbesondere Promotoren. Für die Expression eines erfindungsgemäßen Nucleinsäuremoleküls in prokaryontischen Zellen stehen eine Reihe von Promotoren zur Verfügung, z.B. der E. coli lac- oder trp-Promotor, der P_R - oder P_L -Promotor des Lambda-Phagen, lacI, lacZ, T3, T7, gpt, etc. Eukaryontische Promotoren sind beispielsweise der CMV immediate early-Promotor, der HSV-Promotor, der Thymidinkinase-Promotor, der SV40-Promotor, LTRs von Retroviren und der Maus MetallothioninI-Promotor. Es ist bereits eine Vielzahl von Expressionsvektoren für die Expression in prokaryontischen oder eukaryontischen Zellen beschrieben, z.B. für Eukaryonten pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) or GEM1 (Promega Biotec, Madison, WI, USA), pSV2CAT, pOG44 und für Prokaryonten pQE70, pQE60, pBluescript SK, etc. Neben Promotoren können erfindungsgemäße Vektoren auch Elemente zur weiteren Steigerung der Transkription enthalten, wie z.B. sogenannte Transkriptions-Enhancer. Beispiele dafür sind der SV40-Enhancer, der Polyoma-Enhancer, der Cytomegalovirus early promoter-Enhancer und Adenovirus-Enhancer.



Die vorliegende Erfindung betrifft ferner Wirtszellen, insbesondere prokaryontische oder eukaryontische Wirtszellen, die mit einem erfindungsgemäßen Nucleinsäuremolekül oder Vektor transformiert sind. Beispiele für solche Zellen sind bakterielle Zellen, wie z.B. *E. coli*, *Streptomyces*, *Bacillus*, *Salmonella typhimurium*; Pilzzellen, wie beispielsweise Hefezellen, insbesondere *Saccharomyces cerevisiae*; Insektenzellen, wie z.B. *Drosophila*- oder SF9-Zellen; tierische Zellen, wie z.B. CHO oder COS-Zellen; Pflanzenzellen etc.

Weiterhin betrifft die vorliegende Erfindung ein Verfahren zur Herstellung eines Proteins, das von einem erfindungsgemäßen Nucleinsäuremolekül codiert wird, wobei eine erfindungsgemäße Wirtszelle unter Bedingungen kultiviert wird, die die Expression des Proteins erlauben, und das Protein anschließend aus den Zellen und/oder dem Kulturmedium gewonnen wird. Verfahren zur Expression von Fremdproteinen in verschiedenen Arten von Wirtszellen sowie zur Gewinnung des produzierten Proteins sind dem Fachmann geläufig.

Gegenstand der Erfindung ist ferner ein Protein, das von einem erfindungsgemäßen Nucleinsäuremolekül codiert wird oder das durch ein erfindungsgemäßes Verfahren erhältlich ist.

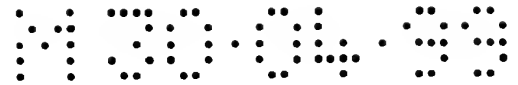
Ferner betrifft die vorliegende Erfindung Antikörper, die gegen die erfindungsgemäßen Proteine gerichtet sind. Vorzugsweise erkennen solche Antikörper spezifisch ein erfindungsgemäßes Protein, d.h. sie zeigen keine nennenswerte Kreuzreaktion mit anderen Proteinen. Der Begriff "Antikörper" umfaßt dabei sowohl monoclonale als auch polyclonale Antikörper, ebenso wie Fragmente von Antikörpern, wobei diese Fragmente ein erfindungsgemäßes Protein erkennen, z.B. Fab-Fragmente. Der Begriff Antikörper umfaßt ebenfalls chimäre Antikörper sowie humanisierte Antikörper. Verfahren zur Herstellung von monoclonalen oder polyclonalen Antikörpern sind dem Fachmann geläufig und sind beschrieben. Zur Her-



stellung von monoclonalen Antikörpern kann beispielsweise die Hybridoma-Technik (Köhler und Milstein, Nature 256 (1975), 495-497), die Trioma-Technik, die menschliche B-Zellhybridom-Technik (Kozbor et al., Immunology Today 4 (1983), 72) oder die EBV-Hybridoma-Technik (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985), 77-96) verwendet werden.

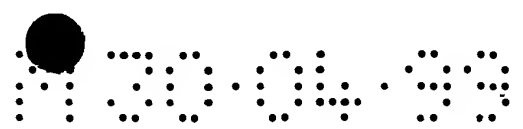
Ferner betrifft die vorliegende Erfindung Nucleinsäuremoleküle von mindestens 15, vorzugsweise mehr als 50 und besonders bevorzugt mehr als 200 Nucleotiden Länge, die spezifisch mit einem Strang eines erfindungsgemäßen Nucleinsäuremoleküls hybridisieren. Spezifisch hybridisieren bedeutet hierbei, daß diese Moleküle mit Nucleinsäuremolekülen hybridisieren, die ein erfindungsgemäßes Protein codieren, jedoch nicht mit Nucleinsäuremolekülen, die andere Proteine codieren. Hybridisieren bedeutet dabei vorzugsweise Hybridisieren unter stringenten Bedingungen (s.o.). Solche Nucleinsäuremoleküle können beispielsweise als Primer für die Amplifikation mittels PCR oder als Hybridisierungsproben verwendet werden. Insbesondere betrifft die Erfindung solche Nucleinsäuremoleküle, die mit Transkripten von erfindungsgemäßen Nucleinsäuremolekülen hybridisieren und dadurch deren Translation verhindern können. Solche Nucleinsäuremoleküle können beispielsweise Bestandteile von antisense-Konstrukten oder Ribozymen sein.

Ferner betrifft die vorliegende Erfindung diagnostische Zusammensetzungen enthaltend ein erfindungsgemäßes Nucleinsäuremolekül oder Vektor, ein erfindungsgemäßes Protein und/oder einen erfindungsgemäßen Antikörper. Die erfindungsgemäßen Nucleinsäuremoleküle können beispielsweise eingesetzt werden, um die Lokalisation des entsprechenden Gens auf einem Chromosom zu bestimmen. Dies kann Aufschluß über die Korrelation mit Genen geben, die mit bestimmten Krankheiten assoziiert sind. eine Methode zur Bestimmung der Lokalisation ist beispielsweise die "Fluorescent in situ hy-



bridisation" (Fish) beschrieben in Verma et al. (Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York (1988)). Weiterhin können erfindungsgemäße Nucleinsäuremoleküle verwendet werden, um festzustellen, ob bestimmte Individuen Mutationen in den entsprechenden Sequenzen aufweisen. Ebenso können Antikörper als Nachweisreagenzien für die Anwesenheit eines erfindungsgemäßen Proteins in einer Probe verwendet werden.

Ein weiterer Gegenstand der vorliegenden Erfindung sind pharmazeutische Zusammensetzungen, die ein erfindungsgemäßes Nucleinsäuremolekül, einen erfindungsgemäßen Vektor, ein erfindungsgemäßes Protein und/oder einen erfindungsgemäßen Antikörper enthalten, gegebenenfalls in Kombination mit einem pharmazeutisch verträglichen Träger. So können beispielsweise erfindungsgemäße Nucleinsäuremoleküle oder Vektoren im Rahmen von Gentherapie eingesetzt werden, um Krankheitszustände zu behandeln, die auf eine Dysfunktion des entsprechenden Gens zurückzuführen sind, beispielsweise auf eine zu geringe oder zu hohe Expression des erfindungsgemäßen Proteins in einem Individuum. Insbesondere können die erfindungsgemäßen Nucleinsäuremoleküle im Zusammenhang mit "gene targeting" und/oder "gene replacement" eingesetzt werden, um ein mutiertes Gen wieder in eine funktionelle Form zu überführen oder um ein mutiertes Gen durch homologe Rekombination zu erzeugen (siehe z.B. Mouellic, Proc. Natl. Acad. Sci. USA 87 (1990), 4712-4716; Joyner, Gene Targeting, A Practical Approach, Oxford University Press). Ebenso kann ein erfindungsgemäßes Protein oder ein erfindungsgemäßer Antikörper eingesetzt werden, um gegebenenfalls die Menge an entsprechendem Protein in einem Individuum zu regulieren. Beispiele für geeignete pharmazeutisch verträgliche Träger sind dem Fachmann geläufig und umfassen beispielsweise phosphatgepufferte Salzlösungen, Wasser, Emulsionen, wie z.B. Öl/Wasser-Emulsionen, sterile Lösungen etc. Zusammensetzungen, die derartige Träger enthalten, können nach gängigen Verfahren formuliert werden. Die pharmazeutischen Zusammen-



setzungen können dem betroffenen Individuum in einer geeigneten Dosis verabreicht werden. Arten der Verabreichung sind beispielsweise intravenös, intraperitoneal, subcutan, intramusculär, topisch oder intradermal. Die Dosierung hängt dabei von vielen Faktoren ab, z.B. von der Größe, dem Geschlecht, dem Gewicht, dem Alter des Patienten, sowie der Art der speziell verabreichten Verbindung, der Art der Administration etc. Im allgemeinen liegt die täglich verabreichte Dosis bei 1 μg bis 10mg Einheiten pro Tag. Im Zusammenhang mit der intravenösen Injektion von DNA sind Dosierungen von 10^6 bis 10^{22} Kopien des DNA-Moleküls gängig. Die Zusammensetzungen können lokal oder systemisch verabreicht werden. Im allgemeinen wird die Verabreichung parenteral erfolgen, z.B. intravenös. DNA kann auch direkt an dem Zielort verabreicht werden, z.B. durch biolistische Verabreichung.

Die vorliegende Erfindung betrifft ferner ein Verfahren zur Herstellung eines transgenen nicht-menschlichen Tiers, vorzugsweise einer transgenen Maus, das die Einführung eines erfindungsgemäßen Nucleinsäuremoleküls oder Vektors in einer Keimzelle embryonale Zelle, eine Eizelle oder eine davon abgeleitete Zelle umfaßt. Das in einem solchen Verfahren als Donor der Zellen verwendete nicht-menschliche Tier kann z.B. ein gesundes nicht-transgenes Tier sein oder ein Tier, das eine Krankheit oder Störung aufweist, insbesondere eines, das eine Wachstumsstörung, vorzugsweise eine die Knochen betreffende Wachstumsstörung aufweist. Eine derartige Krankheit oder Störung kann dabei angeboren oder natürlicherweise entstanden sein oder sie kann durch genetische Manipulation hervorgerufen sein, z.B. durch die Einführung und/oder Expression einer Fremd-DNA.

Gegenstand der vorliegenden Erfindung sind ferner transgene nicht-menschliche Tiere, die mit einem erfindungsgemäßen Nucleinsäuremolekül oder Vektor transformiert sind oder die erhältlich sind durch das oben beschriebene Verfahren. Vorzugsweise ist in solchen transgenen Tieren das erfindungsge-

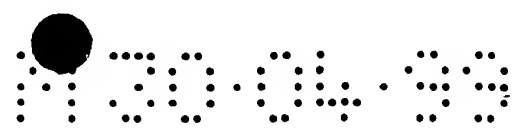


mäße Nucleinsäuremolekül stabil in das Genom integriert. Beispiele für transgene Tiere sind transgene Ratten, Hamster, Hunde, Affen, Kaninchen oder Schweine. Bevorzugt sind transgene Mäuse.

Die vorliegende Erfindung betrifft ebenfalls transgene nicht-menschliche Tiere, insbesondere Mäuse, bei denen die Expression des erfindungsgemäßen Proteins verringert ist. Eine derartige Verringerung kann beispielsweise durch genetische Veränderung der Zellen der Tiere erreicht werden, so daß diese eine antisense-RNA, ein Ribozym oder eine Cosuppressions-RNA exprimieren, die zur Reduktion der Expression erfindungsgemäßer Proteine in den Zellen führt. Alternativ kann eine Verringerung der Expression der erfindungsgemäßen Proteine auch dadurch erreicht werden, daß mindestens eine, vorzugsweise alle Kopien eines einem erfindungsgemäßen Molekül entsprechenden Gens im Genom der Zellen inaktiviert werden. Eine derartige Inaktivierung kann z.B. durch die Insertion von Fremd-DNA in codierende oder nicht codierende Bereiche des entsprechenden Gens erreicht werden. Möglich ist ebenso die Inaktivierung der regulatorischen Regionen des Gens. Möglich ist ferner die Deletion von Bereichen des Gens.

Die vorliegende Erfindung betrifft auch die Möglichkeit, erfindungsgemäße Nucleinsäuremoleküle in vivo, d.h. in Zellen, Zellkulturen oder Organismen zu aktivieren ("Genaktivierung"). Dies kann beispielsweise dadurch erfolgen, daß in das Genom einer Zelle, die ein erfindungsgemäßes Nucleinsäuremolekül enthält, vor das erfindungsgemäße Nucleinsäuremolekül ein Promotor insertiert wird, der beispielsweise konstitutiv ist und eine sehr hohe Expression gewährleistet, oder ein Promotor, der induzierbar ist und bei Induktion eine sehr hohe Expression gewährleistet.

Das im Rahmen der vorliegenden Erfindung hergestellte Plasmid HSL1 wurde bei der als internationale Hinterle-

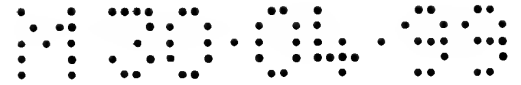


gungsstelle anerkannten Deutschen Sammlung von Mikroorganismen und Zellkulturen (DSMZ) in Braunschweig, Bundesrepublik Deutschland entsprechend den Anforderungen des Budapester Vertrages am 25. März 1998 unter der Hinterlegungsnummer DSM 12073 hinterlegt.

Figur 1 zeigt eine heterozygote LOBO-Maus mit einer Insertion im LOBO-Gen (oben) im Vergleich zu einer Wildtyp-Maus. Die beiden Tiere sind Geschwister und sind ca. 6 Wochen alt.

Figur 2 zeigt die Sequenzierungsstrategie für die Sequenzierung des murinen und menschlichen LOBO-Gens. Da bisher nur das 3'-Ende des Gens sequenziert wurde, wurden die Exons beginnend vom 3'-Ende mit 1,2,3 etc. durchnummeriert. Es wurden drei murine Wildtyp-Cosmid-Clone (Mitte), zwei Plasmid-Clone aus der transgenen LOBO-Maus (oben) und ein humaner P1-Clon (unten) sequenziert. Die Pfeile kennzeichnen die bisher bekannten Exons. Sieben Exons befinden sich auf der genomischen Sequenz, das achte Exon existiert bisher nur auf einem EST-Clon. Die Plasmid-Clone aus der transgenen LOBO-Maus (oben) enthalten das eingeführte künstliche Gen und die angrenzenden murinen Sequenzen. Diese murinen Sequenzen sind mit den entsprechenden Sequenzen der Wildtyp-Maus bis auf 10 Basenpaare identisch, die in der transgenen Maus durch das künstliche Gen ersetzt worden sind.

Figur 3 zeigt einen Sequenzvergleich zwischen dem LOBO-Protein aus Mensch (HS) und Maus (MM) mit verwandten Proteinen aus *S. cerevisiae* (SC-Dis 3), *S. pombe* (Spom 1. prot), *C. elegans* (CE) und *Lactococcus* (Lactococ). Aminosäurereste, die bei mehr als der Hälfte der aufgeführten Organismen an der angegebenen Position vorkommen, sind grau hinterlegt. Es



ist jeweils nur der 3'-Bereich der Proteine dargestellt, da der 5'-Bereich für die Proteine aus *Lactococcus*, Mensch und Maus noch nicht zur Verfügung steht.

Figur 4 zeigt einen histologischen Dünnschnitt durch Knochen der Wildtyp-Maus (A, C, E) und der heterozygoten LOBO-Maus (B, D, F).

Die Wandung der Knochen ist bei LOBO-Mäusen etwa 3 mal so dick wie beim Wildtyp (siehe F im Vergleich zu E). Dennoch sind die Knochen der LOBO-Mäuse brüchiger als die des Wildtyps, da die stabilisierenden Elemente im Knochenmark in Bezug auf Anzahl und Größe signifikant reduziert sind (siehe "blaue Inseln" innerhalb des Knochenmarks, A im Vergleich zu B). Das Loch im Knochenmarksgewebe der LOBO-Maus (B) ist ein präparationsbedingter Artefakt.

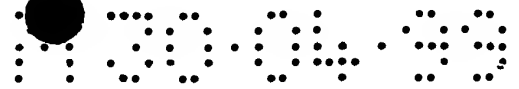
Die folgenden Beispiele veranschaulichen die Erfindung.

Beispiel 1

Auffinden einer Maus mit einem veränderten Knochenwachstum

Im Zusammenhang mit der Untersuchung eines bestimmten künstlichen Proteins wurde eine transgene Maus erzeugt, die als Donor-Maus dienen sollte, d.h. als Spender für das künstliche Protein. Dieses Protein sollte in bestimmten Geweben der "Donor"-Maus exprimiert werden, ohne jedoch in dieser Maus eine Funktion zu haben. Erst nach der Kreuzung der Donor-Maus mit einer geeigneten transgenen Empfänger-Maus sollte das Protein wirksam werden und bestimmte Gene der Empfänger-Maus aktivieren.

Die Herstellung der Donor-Maus erfolgte durch Insertionsmutagenese im Rahmen der Durchführung eines transgenen Mausprojektes. Das eigentliche Ziel des Projektes bestand darin, transgene Mäuse zu etablieren, die den Tetracyclin



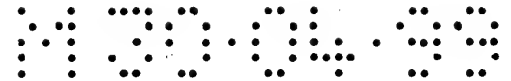
regulierbaren Transaktivator (rTA) in lymphoiden Zellen exprimieren. Die für die Mikroinjektion in Pronuclei verwendete Expressionskassette umfaßte die folgenden Elemente in 5'-3'-Richtung: μ E: Enhancer aus dem Intron der schweren Kette der Immunglobulin-Gene der Maus (700 bp); einen synthetischen Promotor, bestehend aus einem Oktamer-Oligonucleotid und dem Minimal-Promotor des Maus- β -Globin-Gens (Wirth et al., Nature 329 (1987), 174-178) und ein Tet-R/VP16-Konstrukt. Die Enhancer/Promotor-Kombination wurde beschrieben in Annweiler et al. (Nucl. Acids. Res. 20 (1990), 1503-1509). Das Tet-R/VP16-Konstrukt ist beschrieben in Gossen und Bujard (Proc. Natl. Acad. Sci. USA 89 (1992), 5547-5551). Die Gesamtgröße des DNA-Fragments beträgt etwa 3 kb.

Zur Herstellung der transgenen Mäuse wurden 1-2 picoliter einer DNA-Lösung, die die oben beschriebene Expressionskassette enthielt (Konzentration 1 ng/ μ l) in den männlichen Vorkern einer befruchteten Eizelle einer NMRI-Maus injiziert. Anschließend wurde die Eizelle in den Eileiter einer scheinschwangeren weiblichen Ammenmaus transplantiert und von dieser zur Geburt ausgetragen.

Transgene Donor-Mäuse zeigen normalerweise keinen Phänotyp, da das künstliche Gen einfach in befruchtete Eizellen gespritzt wird und sich rein zufällig in irgendeinen Bereich des murinen Genoms integriert.

Da lediglich ca. 5 % des Genoms codierende Bereiche umfassen, ist die Wahrscheinlichkeit, daß ein Defekt in einem essentiellen Gen verursacht wird, entsprechend gering. Darüber hinaus ist das Säugergenom diploid, d.h. alle Gene sind in doppelter Ausführung vorhanden. Da einem möglicherweise mutierten Gen in der Regel eine voll funktionsfähige Kopie gegenübersteht, die den Defekt in der mutierten Version kompensieren kann, sind die meisten Mutationen rezessiv, d.h. sie kommen nicht zur Ausprägung, wenn nur eine Kopie des Gens betroffen ist.

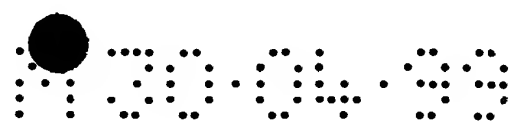
Eines der im Rahmen der Herstellung der oben beschriebenen Donor-Mäuse erhaltene Founder-Tier zeigte nun überraschen-



derweise einen extrem auffälligen Phänotyp insfoern, als es deutlich größer als die im gleichen Wurf geborenen Geschwister war. Auffällig waren der deutlich verlängerte Schwanz sowie die verlängerten Gliedmaßen, insbesondere die langen Zehen. Der Größenunterschied im Vergleich zu den normalen Mäusen verstärkte sich noch signifikant über die nachfolgenden Wochen, dabei bildete sich eine deutliche Skoliose aus. Alle Knochen mit Ausnahme der Schädelknochen sind um das 1,3- bis 1,5-fache verlängert. Infolgedessen ist die transgene Maus insgesamt ca. 1,5 mal länger als eine entsprechende Wildtyp-Maus (siehe Figur 1). Aufgrund der stark verlängerten Knochen ("long bones") wurde die transgene Maus als LOBO-Maus bezeichnet. Dieser Phänotyp wird stabil und autosomal dominant weitervererbt. Bei Kreuzung einer transgenen Mutante mit einer gesunden Wildtyp-Maus zeigen 50 % der Nachkommen den oben beschriebenen Phänotyp. Alle transgenen Nachkommen dieses Founder-Tieres zeigten, unabhängig vom Geschlecht, den gleichen Phänotyp. In der Zwischenzeit wurden mehr als 200 solcher transgenen Tiere charakterisiert. In allen Fällen korrelierte das Vorhandensein des Transgens mit diesem Phänotyp.

Die weitere histologische Analyse von Knochen transgener Mäuse ergab, daß zum einen eine Osteoporose vorliegt als auch eine veränderte innere Knochenstruktur, insbesondere eine deutliche Reduktion der trabekulären Strukturen. Das verstärkte Knochenwachstum läßt sich bereits sehr früh kurz nach der Geburt der Mäuse anhand der verlängerten Gliedmaßen und Zehen feststellen, was darauf hinweist, daß das verstärkte Knochenwachstum bereits von Anbeginn der Knochenentwicklung im Embryo auftritt. Histologisch zeigen sich allerdings in der Knochenwachstumszone keine Auffälligkeiten.

Die Lebenserwartung der mutanten Tiere ist im Vergleich zu ihren Wildtyp-Geschwistern reduziert. Beginnend mit ca. 6 Wochen nach der Geburt zeigen die LOBO-Mäuse eine erhöhte Mortalität, und nach einem knappen Jahr sind alle Mäuse aus derzeit noch nicht bekannten Gründen gestorben. Homozygote Mäuse sind lebensfähig. Obwohl bislang erst zwei Würfe mit



homozygoten Tieren erhalten wurden, werden die homozygoten Tiere der erwarteten Anzahl geboren. Allerdings sind sie deutlich weniger gut entwickelt als Wildtyp-Tiere und heterozygote Tiere. Sie zeigen aber ebenso das verstärkte Knochenwachstum, was an den längeren Fingern eindeutig zu erkennen ist.

Die Analyse der transgenen Tiere ergab insbesondere, daß neben der signifikanten Verlängerung der Knochen noch drei weitere Merkmale für die transgenen Mäuse charakteristisch sind:

1. Die Wandung der Knochen ist bei der LOBO-Maus mindestens doppelt so dick wie beim Wildtyp. Das Knochenmark der normalen Maus enthält stabilisierende Elemente, die bei der transgenen Maus nur in sehr reduzierter Menge vorhanden sind, so daß die Knochen der transgenen Mäuse trotz der verdickten Wandung brüchiger sind als die der Wildtyp Mäuse (siehe Figur 4).
2. Die Lebenserwartung der LOBO-Mäuse ist signifikant verkürzt und beträgt maximal 40 Wochen. Die durchschnittliche Lebenserwartung liegt bei 25 Wochen, während die Wildtyp-Mäuse 1-2 Jahre alt werden.

Beispiel 2

Genetische Analyse der transgenen Maus

Die molekulare Analyse der Ursache der Mutation ergab, daß ca. 1,5 Kopien des Transgens in das Intron eines endogenen Gens insertiert wurden. Expressionsstudien (Northern-Analysen) mit einer cDNA-Probe des endogenen Gens ergab, daß das betroffene Gen offensichtlich ubiquitär exprimiert wird. Während die meisten Organe im Northern nur eine einzelne Bande ergeben (ca. 4 kb) findet sich in der Leber noch ein zusätzliches kürzeres Transkript (ca. 2kb). Es ist zur Zeit unklar, ob dieses kleinere Transkript a) eine Splice-Variante des Gens darstellt, b) auf die Benutzung eines alterna-



tiven Promotors zurückzuführen ist oder c) die Kreuzreaktion mit einem verwandten Gen darstellt. Im Vergleich zu den Wildtyp-Tieren, findet man in den heterozygoten Tieren nur ca. 50 % der mRNA für dieses Gen, wenn eine Probe vom 3'-Bereich der Insertionsstelle verwendet wird.

Beispiel 3

Identifizierung des LOBO-Gens

Um herauszufinden, welches Gen (oder welche Gene) für den LOBO-Phänotyp verantwortlich sind, wurde der mutierte Bereich aus der transgenen Maus in Bakterien subcloniert. Die Lokalisierung des mutierten Bereiches im Genom der Maus und die anschließende Subclonierung waren möglich, weil die Nucleotidsequenz des eingangs erwähnten künstlichen Gens bekannt war und man diese Information in entsprechenden molekularbiologischen Experimenten nutzen konnte. Zur Identifizierung des Gens, das im folgenden "LOBO-Gen" genannt wird, wurden 6 kb aus dem subclonierten Bereich der transgenen Maus sowie 87 kb aus der entsprechenden homologen Region der Wildtyp-Maus sequenziert. Der gesamte sequenzierte Bereich der genomischen DNA-Clone aus Maus ist in Seq ID No. 5 und 6 dargestellt. Der sequenzierte Bereich umfaßt insgesamt 86902 Basenpaare. Aus technischen Gründen wurde dieser Bereich in zwei Bereiche aufgeteilt, wobei die ersten 49999 Basenpaare in Seq ID No. 5 dargestellt sind und ein Exon umfassen und die sich an diesen Bereich am 3'-Ende anschließenden verbleibenden 36901 Basenpaare in Seq ID No. 6 dargestellt sind. Die Exons sind an den folgenden Positionen lokalisiert:

Seq ID No. 5:	8520 - 8753
Seq ID No. 6:	12487 - 12660
	15497 - 15644
	15908 - 16038
	16148 - 16252

17293 - 17394

18083 - 18556

Der offene Leserahmen beginnt dabei an Position 8520 in Seq ID No. 5. Das Stopcodon befindet sich an der Position 18202 in Seq ID No. 6. Der codierende Bereich codiert die in Seq ID No. 2 dargestellte Aminosäuresequenz. Eine detaillierte Computeranalyse der Sequenzdaten führte zur Identifizierung eines Gens, das aus mindestens 8 codierenden Abschnitten ("Exons") besteht und mindestens 60.000 Basen lang ist, wahrscheinlich aber sehr viel länger. Der bisher identifizierte, codierende Bereich, der in Seq ID No. 1 dargestellt ist, trägt die Information für 393 Aminosäuren. Sequenzvergleiche mit Sequenzen bekannter Proteine ergab, daß zu diesem Protein ein verwandtes Protein des Spulwurms *Caenorhabditis elegans* existiert, das im vorderen Bereich um etwa 400 Aminosäuren länger ist als das Protein der Maus. Von daher ist anzunehmen, daß das murine LOBO-Gen noch deutlich länger als 60 kb ist und bisher nur partiell sequenziert wurde.

Aufgrund der murinen Sequenzdaten wurde eine DNA-Sonde konstruiert, mit deren Hilfe ein humaner P1-Clon isoliert wurde, der das menschliche LOBO-homologe Gen trägt. Die bisher erhaltene Sequenz des menschlichen genomischen Clons ist in Seq ID No. 7 dargestellt. Die Exons sind an den folgenden Positionen lokalisiert:

1 - 136
3971 - 4118
4500 - 4630
4762 - 4866
5904 - 6005
6600 - 7109

Das erste Nucleotid des offenen Leserasters liegt an Position 2. Das Stopcodon befindet sich an der Position 6759. Die durch den codierenden Bereich dargestellte Aminosäuresequenz ist in Seq ID No. 4 dargestellt. Ein Clon enthaltend die menschliche genomische Sequenz wurde hinterlegt unter DSM 12073. Die bisher vorliegenden Sequenzdaten zeigen, daß



auch das humane Gen bisher nur partiell cloniert wurde. Eine Übersicht über die erhaltenen und sequenzierten Clone aus Maus und Mensch ist schematisch in Figur 2 dargestellt.

Beispiel 4

Chromosomale Lokalisierung des LOBO-Gens

Einer der erhaltenen Maus-Clone, der einen Teil des murinen LOBO-Gens repräsentiert, wurde mit Hilfe von "Fish" (fluorescent in situ hybridization) farbmarkiert und auf komplette, murine (Metaphase-) Chromosomen hybridisiert. Es resultierte ein Farbsignal in der Bande 1D auf dem Chromosom 1 der Maus. Diese Region ist mit der Bande 2q35 auf dem humanen Chromosom 2 homolog. In diesem Abschnitt befindet sich auch ein Gen für alkaline Phosphatase, dessen Position in der Literatur genau bekannt ist. Dieses Phosphatase-Gen befindet sich sowohl bei der murinen Sequenz, als auch bei der humanen Sequenz ca. 20 kb hinter dem LOBO-Gen, so daß dessen chromosomale Lokalisation in Relation zum Phosphatase-Gen sehr genau angegeben werden kann.

Beispiel 5

Chrakterisierung der Mutation im LOBO-Gen der transgenen Maus

Die weitere genetische Analyse des LOBO-Gens der transgenen Maus zeigte, daß sich die Insertion des künstlichen Gens innerhalb des (nicht-codierenden) Introns zwischen den Exonen 7 und 6 befindet. (Da bisher lediglich das 3'-Ende des Gens sequenziert ist, wurden die Exons ausgehend vom 3'-Ende aus fortlaufend mit 1, 2, 3 etc. in Richtung 5'-Ende numeriert). Die Insertion ist 48,2 kb von Exon 7 und 5,6 kb von Exon 6 entfernt und hat zur Deletion von 10 Basenpaaren geführt. Letzteres sollte im Intron kein Problem bereiten, da dieser Bereich ohnehin nicht codiert. Es ist vielmehr anzunehmen,



daß es aufgrund der künstlichen DNA-Insertion zur Störung bei der Reifung der mRNA kommt. Bei diesem Prozeß werden die Introns, die in der primären RNA noch enthalten sind, herausgeschnitten ("splicing"). Für dieses Herausschneiden sorgen gewisse Sequenzsignale. Solche Signale sind auch in dem künstlich eingeführten Gen enthalten, so daß es vermutlich zu einem sogenannten "aberranten splicing" kommt. Infolgedessen wird die Bildung einer funktionsfähigen mRNA verhindert und das entsprechende Protein kann nicht produziert werden.

Tatsächlich hat die experimentelle Überprüfung der LOBO-Expression (durch "Northern Blot") ergeben, daß in heterozygoten LOBO-Mäusen nur noch etwa die Hälfte an mRNA produziert wird, im Vergleich zur Wildtyp-Maus. In homozygoten LOBO-Mäusen kann überhaupt keine LOBO-mRNA mehr nachgewiesen werden.

Es ist daher anzunehmen, daß die Mutation in der transgenen LOBO-Maus die Genexpression auf post-transkriptionaler Ebene abschaltet. Offenbar sinkt dann bereits in den heterozygoten Mäusen die Menge an produziertem LOBO-Protein unter einen kritischen Schwellenwert, was dann zu dem beobachteten, dominanten Phänotyp führt.

Beispiel 6

Charakterisierung des LOBO-Proteins

Die aus den einzelnen Exons der sequenzierten Bereiche des LOBO-Gens abgeleitete Aminosäuresequenz von Maus und Mensch wurde mit bekannten Proteinen verglichen. Dabei stellte sich heraus, daß das Lobo-Protein keiner der bisher identifizierten Proteinklassen zugeordnet werden kann. Die Aminosäuresequenz des Proteins weist jedoch Bereiche auf, die extrem konserviert sind sowohl zwischen Mensch und Maus, als auch zu Proteinen aus *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, einem Bakterium der Gattung *Leuconostoc* und aus *Caenorhabditis elegans* (siehe Figur 3). Von dem Protein aus

Hefe (Dis3) ist bekannt, daß es eine Rolle in der Zellzykluskontrolle spielt. Die Funktion der anderen Proteine ist bisher nicht untersucht.

SEQUENZPROTOKOLL



(1) ALLGEMEINE ANGABEN:

(i) ANMELDER:

- (A) NAME: André Rosenthal
- (B) STRASSE: Beutenbergstr. 11
- (C) ORT: Jena
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 07745

(ii) BEZEICHNUNG DER ERFINDUNG: Nucleinsaeuremolekuele codierend Proteine, die die Knochenentwicklung beeinflussen

(iii) ANZAHL DER SEQUENZEN: 7

(iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

(2) ANGABEN ZU SEQ ID NO: 1:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1550 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

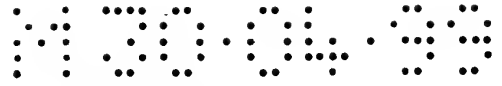
- (A) ORGANISMUS: Mus musculus

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE: 2..1180

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

C CTC GGC CGA AGT AAA GTA GCT GCT GAG AGA GCC ACA AGT GTC TAC	46
Leu Gly Arg Ser Lys Val Ala Ala Glu Arg Ala Thr Ser Val Tyr	
1 5 10 15	
TTG GTC CAG AAG GTG GTC CCC ATG CTT CCC AGG CTT CTG TGT GAG GAA	94
Leu Val Gln Lys Val Val Pro Met Leu Pro Arg Leu Leu Cys Glu Glu	
20 25 30	
CTC TGC AGC CTC AAC CCC ATG ACT GAC AAG CTG ACC TTC TCT GTG ATC	142
Leu Cys Ser Leu Asn Pro Met Thr Asp Lys Leu Thr Phe Ser Val Ile	
35 40 45	
TGG AAG CTG ACC CCT GAA GGC AAG ATC CTT GAA GAG TGG TTT GGC CGC	190
Trp Lys Leu Thr Pro Glu Gly Lys Ile Leu Glu Glu Trp Phe Gly Arg	
50 55 60	



ACT	ATC	ATC	CGT	TCT	TGC	ACC	AAA	CTG	AGC	TAC	GAC	CAT	GCC	CAG	AGC	238
Thr	Ile	Ile	Arg	Ser	Cys	Thr	Lys	Leu	Ser	Tyr	Asp	His	Ala	Gln	Ser	
65						70					75					
ATG	ATC	GAA	AAT	CCA	ACT	GAG	AAG	ATC	CCT	GAG	GAA	GAG	CTT	CCC	CCA	286
Met	Ile	Glu	Asn	Pro	Thr	Glu	Lys	Ile	Pro	Glu	Glu	Glu	Leu	Pro	Pro	
80					85				90						95	
ATT	TCT	CCA	GAG	CAC	AGC	GTC	GAG	GAG	GTG	CAC	CAG	GCA	GTC	CTG	AAC	334
Ile	Ser	Pro	Glu	His	Ser	Val	Glu	Glu	Val	His	Gln	Ala	Val	Leu	Asn	
				100					105					110		
CTG	CAC	AGC	ATT	GCA	AAG	CAA	CTC	CGC	CGC	CAG	CGC	TTT	GTA	GAT	GGC	382
Leu	His	Ser	Ile	Ala	Lys	Gln	Leu	Arg	Arg	Gln	Arg	Phe	Val	Asp	Gly	
			115				120						125			
GCA	CTC	CGT	TTA	GAT	CAG	GAG	TTC	ATG	CTC	CTG	GCC	AAC	ATG	GCG	GTG	430
Ala	Leu	Arg	Leu	Asp	Gln	Glu	Phe	Met	Leu	Leu	Ala	Asn	Met	Ala	Val	
		130					135					140				
GCC	CAC	AAG	ATC	TTC	CGC	ACC	TTC	CCT	GAG	CAG	GCC	CTG	CTG	CGC	CGG	478
Ala	His	Lys	Ile	Phe	Arg	Thr	Phe	Pro	Glu	Gln	Ala	Leu	Leu	Arg	Arg	
		145				150					155					
CAT	CCC	CCA	CCA	CAG	ACG	AAG	ATG	CTC	AGT	GAC	CTG	GTG	GAG	TTC	TGT	526
His	Pro	Pro	Pro	Gln	Thr	Lys	Met	Leu	Ser	Asp	Leu	Val	Glu	Phe	Cys	
160					165					170					175	
GAC	CAG	ATG	GGG	CTG	CCC	ATG	GAT	GTC	AGC	TCT	GCA	GGG	GCC	CTA	AAT	574
Asp	Gln	Met	Gly	Leu	Pro	Met	Asp	Val	Ser	Ser	Ala	Gly	Ala	Leu	Asn	
				180					185					190		
ATG	GCA	CTG	TAC	TTC	TGC	TCT	GGG	ATG	CTG	CAG	GAC	CAG	GAG	CAG	TTC	622
Met	Ala	Leu	Tyr	Phe	Cys	Ser	Gly	Met	Leu	Gln	Asp	Gln	Glu	Gln	Phe	
			195				200						205			
CGG	CAT	TAT	GCT	CTC	AAC	GTT	CCC	CTC	TAC	ACA	CAC	TTC	ACC	TCT	CCC	670
Arg	His	Tyr	Ala	Leu	Asn	Val	Pro	Leu	Tyr	Thr	His	Phe	Thr	Ser	Pro	
		210					215					220				
ATC	CGC	CGC	TTT	GCT	GAC	GTC	ATA	GTG	CAC	CGC	CTC	CTG	GCT	GCT	GCT	718
Ile	Arg	Arg	Phe	Ala	Asp	Val	Ile	Val	His	Arg	Leu	Leu	Ala	Ala	Ala	
		225				230					235					
CTG	GGC	TAC	AGT	GAA	CAG	CCA	GAT	GTG	GAG	CCT	GAT	ACC	CTA	CAG	AAG	766
Leu	Gly	Tyr	Ser	Glu	Gln	Pro	Asp	Val	Glu	Pro	Asp	Thr	Leu	Gln	Lys	
240					245					250					255	
CAA	GCT	GAC	CAC	TGC	AAT	GAC	CGT	CGC	ATG	GCT	TCC	AAA	CGT	GTG	CAG	814
Gln	Ala	Asp	His	Cys	Asn	Asp	Arg	Arg	Met	Ala	Ser	Lys	Arg	Val	Gln	
				260					265					270		
GAG	CTC	AGC	ATC	GGC	CTC	TTC	TTC	GCA	GTT	CTA	GTA	AAG	GAG	AGT	GGC	862
Glu	Leu	Ser	Ile	Gly	Leu	Phe	Phe	Ala	Val	Leu	Val	Lys	Glu	Ser	Gly	
			275					280					285			
CCC	CTG	GAG	TCC	GAA	GCC	ATG	GTG	ATG	GGT	GTC	CTG	AAC	CAA	GCT	TTC	910
Pro	Leu	Glu	Ser	Glu	Ala	Met	Val	Met	Gly	Val	Leu	Asn	Gln	Ala	Phe	
		290					295					300				
GAC	GTG	CTG	GTG	CTG	CGC	TTT	GGG	GTG	CAG	AAG	CGC	ATC	TAC	TGC	AAT	958
Asp	Val	Leu	Val	Leu	Arg	Phe	Gly	Val	Gln	Lys	Arg	Ile	Tyr	Cys	Asn	
		305				310					315					
GCA	CTG	GCC	CTG	CGA	TCC	TAC	AGC	TTC	CAG	AAG	GTG	GGG	AAG	AAG	CCA	1006



(A) LÄNGE: 393 Aminosäuren
(B) ART: Aminosäure
(D) TOPOLOGIE: linear

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Leu 1	Gly	Arg	Ser	Lys 5	Val	Ala	Ala	Glu	Arg 10	Ala	Thr	Ser	Val	Tyr 15	Leu
Val	Gln	Lys	Val 20	Val	Pro	Met	Leu	Pro 25	Arg	Leu	Leu	Cys	Glu 30	Glu	Leu
Cys	Ser	Leu 35	Asn	Pro	Met	Thr	Asp 40	Lys	Leu	Thr	Phe	Ser 45	Val	Ile	Trp
Lys	Leu 50	Thr	Pro	Glu	Gly	Lys 55	Ile	Leu	Glu	Glu	Trp 60	Phe	Gly	Arg	Thr
Ile 65	Ile	Arg	Ser	Cys	Thr 70	Lys	Leu	Ser	Tyr	Asp 75	His	Ala	Gln	Ser	Met 80
Ile	Glu	Asn	Pro	Thr 85	Glu	Lys	Ile	Pro	Glu 90	Glu	Glu	Leu	Pro	Pro 95	Ile
Ser	Pro	Glu	His 100	Ser	Val	Glu	Glu	Val 105	His	Gln	Ala	Val	Leu 110	Asn	Leu
His	Ser	Ile 115	Ala	Lys	Gln	Leu	Arg 120	Arg	Gln	Arg	Phe	Val 125	Asp	Gly	Ala

Leu Arg Leu Asp Gln Glu Phe Met Leu Leu Ala Asn Met Ala Val Ala
 130 135 140
 His Lys Ile Phe Arg Thr Phe Pro Glu Gln Ala Leu Leu Arg Arg His
 145 150 155 160
 Pro Pro Pro Gln Thr Lys Met Leu Ser Asp Leu Val Glu Phe Cys Asp
 165 170 175
 Gln Met Gly Leu Pro Met Asp Val Ser Ser Ala Gly Ala Leu Asn Met
 180 185 190
 Ala Leu Tyr Phe Cys Ser Gly Met Leu Gln Asp Gln Glu Gln Phe Arg
 195 200 205
 His Tyr Ala Leu Asn Val Pro Leu Tyr Thr His Phe Thr Ser Pro Ile
 210 215 220
 Arg Arg Phe Ala Asp Val Ile Val His Arg Leu Leu Ala Ala Ala Leu
 225 230 235 240
 Gly Tyr Ser Glu Gln Pro Asp Val Glu Pro Asp Thr Leu Gln Lys Gln
 245 250 255
 Ala Asp His Cys Asn Asp Arg Arg Met Ala Ser Lys Arg Val Gln Glu
 260 265 270
 Leu Ser Ile Gly Leu Phe Phe Ala Val Leu Val Lys Glu Ser Gly Pro
 275 280 285
 Leu Glu Ser Glu Ala Met Val Met Gly Val Leu Asn Gln Ala Phe Asp
 290 295 300
 Val Leu Val Leu Arg Phe Gly Val Gln Lys Arg Ile Tyr Cys Asn Ala
 305 310 315 320
 Leu Ala Leu Arg Ser Tyr Ser Phe Gln Lys Val Gly Lys Lys Pro Glu
 325 330 335
 Leu Thr Leu Val Trp Glu Pro Asp Asp Leu Glu Glu Glu Pro Thr Gln
 340 345 350
 Gln Val Ile Thr Ile Phe Ser Leu Val Asp Val Val Leu Gln Ala Glu
 355 360 365
 Ala Thr Ala Leu Lys Tyr Ser Ala Ile Leu Lys Arg Pro Gly Leu Glu
 370 375 380
 Lys Ala Ser Asp Glu Glu Pro Glu Asp
 385 390

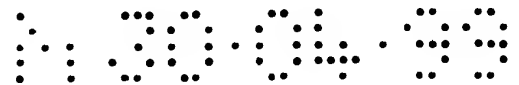
(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1140 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA

(iii) HYPOTHETISCH: NEIN



(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: Homo sapiens

(ix) MERKMAL:

(A) NAME/SCHLÜSSEL: CDS

(B) LAGE:2..782

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

G ATC CAC CGC GCC TTC CCC GAG CAG GCC CTG CTG CGC CGG CAC CCC	46
Ile His Arg Ala Phe Pro Glu Gln Ala Leu Leu Arg Arg His Pro	
365 370 375	
CCG CCC CAA ACA AGG ATG CTC AGT GAC CTG GTG GAA TTC TGC GAC CAG	94
Pro Pro Gln Thr Arg Met Leu Ser Asp Leu Val Glu Phe Cys Asp Gln	
380 385 390	
ATG GGG CTG CCC GTG GAC TTC AGC TCC GCA GGA GCC CTC AAT ATG GCA	142
Met Gly Leu Pro Val Asp Phe Ser Ser Ala Gly Ala Leu Asn Met Ala	
395 400 405	
CTG TAC TTC TGC TCG GGG CTG CTG CAG GAC CCA GCG CAG TTC CGG CAC	190
Leu Tyr Phe Cys Ser Gly Leu Leu Gln Asp Pro Ala Gln Phe Arg His	
410 415 420	
TAC GCG CTC AAT GTG CCC CTG TAC ACA CAC TTC ACC TCG CCC ATC CGC	238
Tyr Ala Leu Asn Val Pro Leu Tyr Thr His Phe Thr Ser Pro Ile Arg	
425 430 435	
CGC TTT GCC GAC GTC CTG GTG CAC CGC CTC CTG GCT GCC GCG TTA GGC	286
Arg Phe Ala Asp Val Leu Val His Arg Leu Leu Ala Ala Ala Leu Gly	
440 445 450 455	
TAT AGG GAG CGA CTA GAC ATG GCG CCC GAT ACC CTG CAG AAA CAG GCG	334
Tyr Arg Glu Arg Leu Asp Met Ala Pro Asp Thr Leu Gln Lys Gln Ala	
460 465 470	
GAC CAC TGT AAC GAC CGC CGC ATG GCG TCC AAG CGC GTG CAG GAG CTC	382
Asp His Cys Asn Asp Arg Arg Met Ala Ser Lys Arg Val Gln Glu Leu	
475 480 485	
AGT ACC AGT CTC TTC TTT GCT GTT CTG GTC AAG GAG AGT GGC CCC CTG	430
Ser Thr Ser Leu Phe Phe Ala Val Leu Val Lys Glu Ser Gly Pro Leu	
490 495 500	
GAG TCA GAA GCC ATG GTG ATG GGC ATC CTG AAG CAA GCC TTC GAC GTG	478
Glu Ser Glu Ala Met Val Met Gly Ile Leu Lys Gln Ala Phe Asp Val	
505 510 515	
CTG GTG CTG CGC TAC GGC GTG CAG AAG CGC ATC TAC TGC AAC GCA CTG	526
Leu Val Leu Arg Tyr Gly Val Gln Lys Arg Ile Tyr Cys Asn Ala Leu	
520 525 530 535	
GCC CTG CGG TCC CAC CAC TTC CAG AAG GTG GGC AAG AAG CCG GAA CTC	574

Ala Leu Arg Ser His His Phe Gln Lys Val Gly Lys Lys Pro Glu Leu
540 545 550

ACG CTG GTC TGG GAG CCT GAG GAC ATG GAG CAG GAG CCA GCA CAG CAG 622
Thr Leu Val Trp Glu Pro Glu Asp Met Glu Gln Glu Pro Ala Gln Gln
555 560 565

GTC ATC ACC ATC TTC AGC CTG GTG GAG GTG GTC CTG CAG GCA GAG TCC 670
Val Ile Thr Ile Phe Ser Leu Val Glu Val Val Leu Gln Ala Glu Ser
570 575 580

ACA GCC CTC AAG TAC AGC GCC ATC CTG AAG CGG CCA GGC ACC CAG GGC 718
Thr Ala Leu Lys Tyr Ser Ala Ile Leu Lys Arg Pro Gly Thr Gln Gly
585 590 595

CAC CTG GGC CCT GAG AAG GAG GAG GAG GAG TCT GAC GGT GAG CCC GAG 766
His Leu Gly Pro Glu Lys Glu Glu Glu Glu Ser Asp Gly Glu Pro Glu
600 605 610 615

GAC TCA AGC ACC AGC T GAGCTCCACC AGCCGCCTGC CCCGCCTGCC 812
Asp Ser Ser Thr Ser
620

CCGCCTGCCT GTCCCGCCAC ACTGGCTTTA GGACCTGTTG ACACGGAGGG GGGTTTTTAA 872

TTTGGTTTTT AACAACTCAG GGGTTTGTTT TTATTTTAT TTAATTTTGT CAGCTCAACT 932

TTTAAACAAA CTGCAGGGGA GAGGGTGGGG CTGGAAGGAA GGCTGAGGCC TGGTCAGCAG 992

TGACCCCAGC AGAGCAGGCC CCAGTCCTCC TGGGAGGCTG GCCCCCTTT TTTCTGGGCC 1052

CTACTGCCCT CCTCTGCCCA GGAAATGGGG GGGTTTCAGC AACTCAGTGT CACAGAATAA 1112

AATCAAGTGT GGAGTGCCAT AAAAAAAA 1140

(2). ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 260 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Ile His Arg Ala Phe Pro Glu Gln Ala Leu Leu Arg Arg His Pro Pro
1 5 10 15

Pro Gln Thr Arg Met Leu Ser Asp Leu Val Glu Phe Cys Asp Gln Met
20 25 30

Gly Leu Pro Val Asp Phe Ser Ser Ala Gly Ala Leu Asn Met Ala Leu
35 40 45

Tyr Phe Cys Ser Gly Leu Leu Gln Asp Pro Ala Gln Phe Arg His Tyr
50 55 60

Ala Leu Asn Val Pro Leu Tyr Thr His Phe Thr Ser Pro Ile Arg Arg
65 70 75 80

Phe Ala Asp Val Leu Val His Arg Leu Leu Ala Ala Ala Leu Gly Tyr
85 90 95

Arg Glu Arg Leu Asp Met Ala Pro Asp Thr Leu Gln Lys Gln Ala Asp
100 105 110

His Cys Asn Asp Arg Arg Met Ala Ser Lys Arg Val Gln Glu Leu Ser
115 120 125

Thr Ser Leu Phe Phe Ala Val Leu Val Lys Glu Ser Gly Pro Leu Glu
130 135 140

Ser Glu Ala Met Val Met Gly Ile Leu Lys Gln Ala Phe Asp Val Leu
145 150 155 160

Val Leu Arg Tyr Gly Val Gln Lys Arg Ile Tyr Cys Asn Ala Leu Ala
165 170 175

Leu Arg Ser His His Phe Gln Lys Val Gly Lys Lys Pro Glu Leu Thr
180 185 190

Leu Val Trp Glu Pro Glu Asp Met Glu Gln Glu Pro Ala Gln Gln Val
195 200 205

Ile Thr Ile Phe Ser Leu Val Glu Val Val Leu Gln Ala Glu Ser Thr
210 215 220

Ala Leu Lys Tyr Ser Ala Ile Leu Lys Arg Pro Gly Thr Gln Gly His
225 230 235 240

Leu Gly Pro Glu Lys Glu Glu Glu Glu Ser Asp Gly Glu Pro Glu Asp
245 250 255

Ser Ser Thr Ser
260

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 49999 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

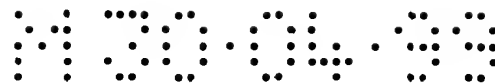
(ii) ART DES MOLEKÜLS: Genom-DNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Mus musculus



(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

GATCAAGTCC AGAACCTCAC ACTGAAACCC AAGCCTTG TG ATGTTCTTAG TGGTGACATT	60
CTTATTCACG TAGTAAATAT TGAATGGTAT TTGTTGCACT CAGATACCAT ACAAGGTATT	120
GAAAATCTCA GACATTTCCC CATCCAGACA GAAGTCCATC TTTCTAGTT GTAGTTGTCT	180
ATTCTCCCTT TCCCCTGGCT GCATGTTTTA AATTTCTTAC AGTAAAGGCA TATTGCAACT	240
TAAAAGCAAA AGTCATTTTG AGACATTTTC GCCTGTTTTT TAATAAGTAG ATGAGATATT	300
GGAGTGCATT TGTAGGCTGA GTGAAAGACA GACAAAGTGA GGAAGGAGTC ACAGTTTGGG	360
AGCCTGGTAA AGAAGGACTC AGCCTATGAG AGCAATGAGT TCCCACAGGA CAAGGGTCAG	420
CTCTTCTCCT ACCTTGACTA GAATAAAGGG AGGGGCTGGG AATGGGGCTC AGTAGACCAT	480
GGGAAGGTGA TTCGATGCTC CCTGTCAGGT TCCCAGGGG TAAATGTCAT TTTCCCTGCA	540
CTCCAGGGCC AGTTCTGTTC CATTCTGTTC TCCTGCCAGA CTCTTTTTTT TTTTTTTACA	600
GTTTTTTTTA ATTAGGTATT TTCTTCATTT ACATTTCAAA TGCTATCCCA AAAGACCCCC	660
CATACCCTCC CCCCCATTCC CCTACCCACC CACTCCCACT TCTTGGCCCT GGTGTTCCCT	720
TGTA CTGGGG CATATAAAGT TTGCAAGACC TATGGGCCTC TCTTCCCAAT GATGGCCGAC	780
TAGGTCATCT TCTGAAACAT ATGCAGCTAG AGACACGAGC TCTGGAGGTA CTGGTTAGTT	840
CATATTGTTG TTCCACCTAT AGGGTTGCAG ACACCTTTAG CTCCTTGAGT ACTTCCTCTA	900
GCTCCTCCAT TGGGGGCCCT GTGTTCCATC CAATAGCTGA CTATGAGCAT CCACTTCTGT	960
GTTTGCCAGG CATCGCATAG CCTCACAAGA GACAGCTGTA TCAGGGTCCT TTCAGCAAAA	1020
TCTTGCTGGT GTATGCAATG GTGTCAGCAT TTGGAGGCTG ATTATGGGAT GGATCCCCGG	1080
GTATTCCTGC CAGACTCTTA AGCCCGGACC AGAGTTTAC GTCTTCCTCA TAGTTCAGTG	1140
CCCTCTACCC AGAAAACACT TTGCCTTGGT TTCACTGTT CTGTTTATTC CTGTTGCTTA	1200
GTGAGATGGT GGGCCCCAAA TAAGCATGTG CATCCCCAGC AGCCACCCCA ATCCTATGAA	1260
CTTGCA TGCT GGGAGTTGTG GAGTGTCTCA GGTAGCCCTG CCATGCTTCC CCACAGAGCT	1320
GCTCTTCATT TCCTTAATGA CCCCTGTGGA CTTTCATACC ATTAACCTGC CAGATGCCAC	1380
CACTGAAAAG CTTGTATTCT TCCTGGGCTA CTGTGGTCCA AAGCAAGACT CCCACAGTGC	1440
CATGTAGCTT AAGGCTTTTC CTAAAAGCAG TGCTAGGTGC TGTGTTTCAT ACCTAGGCAC	1500
CCTACTAAAT ACCTGAGAAA CTCCAGGAGG AAGTAGCTTC AAAGCCTAGT TCTGAGAATC	1560
AGAAATTGTT CCCATAATCT CTCCTCTTAG TCACTACAAG GGCAGAGCC TAGCTGTTTT	1620
ATTCAGGAC TGTCGGTGGG ACCTCTGTAG CAAGGGAGGG ATGGAAGGAG CTGCTGTTCC	1680

ATATCCCTCA AGTCCCAGTT TTCCACTGAA GACACCAGCC AGCTAGATGG .CTTCCCTAAG 1740
 GTCACATCAG AGGAGCAACG GAACTCAGTT GTGAAGCAGT GAAGCTTGAG GATGAAAAGC 1800
 AGAATCCAAA ATGAAACATT TTCAAGATAT GAAATGAGGT GTTTGTTCAT GTAAGCAGCA 1860
 GAAAAGGTTA TGGTGTGGAG TGTCTTTTCA AGGACAAGGG GCTTTATGAG CTGGCTTACA 1920
 ATGGACCTGT TCAAAGGAAG GCTGGGGTAC TAGGTTCAAC AGGCAGAAGG TATCTGTGAT 1980
 GTTTCCTGGA TCCAGAATTC CCCCACCCCC CACCCCCACT GCTACTTCCC ACATTCTCCT 2040
 TCTTCTCCC TCCCCTCCTC CAGTTTCCTT TCTGTACAGA GAGATGAGTC CCAAACATGA 2100
 GCCTTTAATG GGGGACTTTT GGGATAGCAC TGGAAATGTA AACGAGGAAA ATACCTAATA 2160
 AAAAATATTT AAAAAAAAAA GATGCCTCCT GCCAGTCTTG AGGACAGTGG AACACTTTGA 2220
 AGATTATACC TGCTTGAGTA CCTTTACCCA CTGTTACGGG AACACAATTC CTATCTCCTG 2280
 GCCACAGCTA GAGTTTCGGC TCCCTCTAGC CCAATGGTTC TCAGCCTTCC TGATGCTGCA 2340
 ACCCTTTGAT ACAGTTCCTC ATGTTGAGGT GACCCCCAAC CATAAAATTA TTTCATAGCT 2400
 ACTTCATAAC TATAACTTTG GTGCTGTTAT AAACCCTAAT GTTAGCAACC AACATACAGG 2460
 ATGTCTGATA TAATCCCCAA GGGGTTGCAA CCCACAGATT GAAAACCCCT GATCTAGATG 2520
 CTGTATGTGG CAAAGATTTG GTTTCCTCTG CTTCTTGTG TTTGGTTTAG AAGCTTACAT 2580
 AGCTGTCATC AGATCAGGAT GGGAAAGGAC CTAATCTCTC TTGAGACTGA AGGACAAGCC 2640
 AGTGAGTGAT AAGATTGTAT AGTTAATTCC AGCTTCTTCT CTATGCAGAC TCTACCATGT 2700
 GCACAAACTG ACTTAGAACC CAAACAGGCT GGCTAACTTG GAACCAGCCA ACCTGTGTTG 2760
 CTGGGCTTCT AAGGCACTGG TCCTTTCCCA GCCACTGGTG GTCTTGACAC AGCAAGAGCA 2820
 AGCCTGTGAG ATGAAAGGAG CTGCTGCTGG TGGGAGGCAG CCTTGCCACA GTTTCATTCT 2880
 GCCCTGCTGT CTTTCTCTTG TTGTCAGTCT CATTCTGTCA CCTCAGGCCT CAGTTGAGAG 2940
 AGGGCCTAAT GAAGGAGGAC CCCCACCCCT GCCCCCTGCT TATATGAAGC CACCCCATAG 3000
 TTTCTGACTA GTTAGTCACA GGTCAATTCCA TAAGGAATCA GCTTTCCTTC CATCAAGCAA 3060
 CCTCCTGCCC TTTGCTGTCC CCGCCTCTCC ACCTCTGCCC AAGTCATTTT CAGACACTTT 3120
 GTTCTTGACA CCTTTTACTG TCCTTTTGGC CAGGATGGCT GGGATGGCCA GGACGGCCAT 3180
 GTTGGCTGGG ATAGCCATGT TGACCAGACT AGCCTTGCCT TCATAGCTTT AAGAAGCAGC 3240
 AGCAATCTGC TGCCCCCAGG CACCACCACC ACTCCAGACA GCCTGCTTTT GTCCAGTCA 3300
 GGAAAGTGCT TCTTCTGCC TTCCAGGCTT TTTGAACTAA AAGTTCTGTA TGAGGAAGCC 3360
 CAGAGGTTCA GAACTCATTT CACATCTAGT TATTTAAAAT TTAAAATTAG CTCTATTAGT 3420

AGTTTTTTGA ACCAAATATG TCTCAATGAG TTAATATTTT TCAGAGAAAT ATTTTAAAAA 3480
 AGTTCATGGA ATAGGACGGA GGTCCAAAGG TTTCTTCACG CCTTTATATC TATAAATTGT 3540
 AGAAATGAGG TATAATTGTA GAAATATATT TGAGGTATAT TTTGATTCTC ATCATCTACG 3600
 TTACGACCTT CCGCTAAAAA GAAAAAAGTG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG 3660
 TGTGTGTGTG TGTGCGCGCA CACACACATC CACATGAATC CACTATATAT ATATATATTT 3720
 TTTTTACTCT GAACCTTCAG GTATGGACCT AAGAGTTTGC ATGATTCTTG AGTATTTCCC 3780
 ACCTGATTGC CCAGCTTCCC CTGGTGTGTC AAAGTGATGC TCAAAGGCTG TGTACCTGAG 3840
 GCTGGGACCA GCAGCACTGA GTAGGTCAGG AGGGGATACC TCCTTAGATA ATGGGTTTCT 3900
 CAGCCATGTG TCTTCAGTCT GTGGAGAGAC TGTGCTTAAG CTGACATTCT GAACAGTGGC 3960
 ACCCCACAGT ATGTGCTAGA ATCCTGTGTA GAGTTCAGTG TGGCCTGAAT CCTGTGGTTA 4020
 TGCAAAGGAG GCAGGACACG ATCTCCTCAG GGGTACTGTC CATGTGTTCC CTCCTCCTTT 4080
 TTTTTTTCTA CCTTTTCCAT GAAAAGCCCT TTGTCTTCTG CCACTGGCTC TGGTTATGGA 4140
 CTTGGTGTTG ATGTGAGTAC AGTTTTCAGA TTGGAAATTA ATGAGGTGTT CCATTGAGAG 4200
 AAGCCTGACT TCTACCCTGG CTGGCTGCTC CCAGGTTTCC TCCATGTGGG TCTTTGCTGC 4260
 TTTCTCTGTG GGCAGCTGCC CTTGGCTGGC ATTCTTCTAT TGGCTTTCCC CAGAGGTACT 4320
 TTCAAGACTG CTTTCCCAGG CTAGAAACTA TTCTAGTACA TGTCAGCTGT GCCTCCCACA 4380
 AGTCCCAAGC CATGGTAAAG CCAGACAGCC TTGGCTGAGA AGGGAAGTTC GAAAAGGCTC 4440
 TCCTTTGTAT GTTTGTGAAG AAGGGATGAA GGGCAAAAGA GGAAGGGAAA TCAGGTAAAG 4500
 ATGCTATGGA AACCAGCACC TAAAGTAGAA AGTTTGGTAG TGTCCATGTG GGCATTGGAG 4560
 AAAGGCTGTC TTGACAAGAA GGAAACAAAG AAGCAGAGGT ACCTATTAGG TAGAACAGGT 4620
 GCTTCTAATA AGATAGTGTA CTATTAGTAG GCATGTAGCC AGGCTCTGGT GAGGAATAGT 4680
 AGGCAACATA GGGTGACACA TGGCTGCTAG TCAGGGCTCA ACAATCAGAG GGGACTAAGG 4740
 AAGCAACTGA TGTGTAGAGC CAAGACATGT GGGCATGTAG GCAGAAGAAC ATCTAAGAGC 4800
 TTTGTACAGC TTAAGTGTAA GGTGTGTGCA TAAACTTAG AATGCTCTGA GCACTCATCA 4860
 GATTCTACAG CTGTTCTTGC TCCAACCTTG TACAGCAGAA ATCTGCTAAT TGTGTAGTAG 4920
 TTACCTTCAC TTGAGTGTCA TGTACTAGGA AGGAGGATGC AGGCCACAGG AGGACAGATA 4980
 TCAAGACCTG AGTGTGGGGA GGAGTTCATG AGCTAGCTCA CTGGGAGGTG TAGGAATGAA 5040
 AAGGGTGGCA CACAATGTAA GCTGCCACCA TCTGTCAGCA GGCTGAAAAC AGACTGCCTA 5100
 ACACACATGT ACACAGGACT GAGCTGAGGG AGAACTCATT TGGGAAGAAA ATTAAGAAAA 5160

GAAAGAAGCA TAGTGTCCAC ACTTCAGTCT TCATTTTTCT TGAGTTTCAT GTGTTTAGGA 5220
 AATTGTATCT TATATCTTGG GTATCCTAGG TTTTGGGCTA ATATCCACTT ATCAGTGAGT 5280
 ACATATTGTG TGAGTTCCTT TGTGAATGTG TTACCTCACT CAGGATGATG CCCTCCAGGT 5340
 CCATCCATTT GGCTAGGAAT TTCATAAATT CATTCTTTTT AATAGCTGAG TAGTACTCCA 5400
 TTGTGTAGAT GTACCACATT TTCTGTATCC ATTCTCTGTG TGAGGGGCAT CTGGGTCTCT 5460
 TCCAGCTTCT GGCTATTATA AATAAGGCTG CTATGAACAT AGTGGAGCAT GTGTCCTTCT 5520
 TACCAGTTGG GGCATCTTCT GGATATATGC CCAGGAGAGG TATTGCTGGA TCCTCCGGTA 5580
 GTACTATGTC CAATTTTCTG AGGAACCGCC AGACGGATTT CCAGAGTGGT TGTACAAGCC 5640
 TGCAATCCCA CCAACAATGG AGGAGTGTTT CTATTTCTCC ACATCCACGC CAGCATCTGC 5700
 TGTCACCTGA ATTTTGTATC TTAGACATTC TGAAGTAGTG GAGGTGGAAT CTCAGGGTTG 5760
 TTTTGATTTG CATTTCCCTG ATGATTAAGG ATGTTGAACA TTTTTCAGG TGCTTCTCTG 5820
 CCATTCGGTA TTCCTCAGGT GAGAATTCTT TGTTCAGTTC TGAGCCCCAT TTTTAAATGG 5880
 GGTTATTTGA TTTTCTGAAG TCCACCTTCT TGAGTCTTTT ATATATGTTG GATATTAGTC 5940
 CTCTATCTAA TTTAGGATAG GTAAAGATCC TTTCCCAATC TGTGGTGGT CTCTTTGTCT 6000
 TATTGACGGT GTCTTTTGCC TTGCAGAAAC TTTGGAGTTT CATTAGGTCC CATTTGTCAA 6060
 TTCTCGATCT TACAGCACAA GCCATTGCTG TTCTGTTTCTG GAATTTTCTC CCTGTGCCCCA 6120
 TATCTTCAAG GCTTTTCCCC ACTTTCTCCT CTATAAGTTT CAGTGTCTCT GGTTTTATGT 6180
 GAAGTTCTTT GATCCATTTA GATTTGACCT AGTGTGGACA CTATGCCCCCT CCTTAGAAGT 6240
 GGGAAACAAA CACCCTTGGA AGGAGTTACA GAGACAAAGT TTGGAGCTGA GATGAAAGGA 6300
 TGGACCATGT AGAGACTGCC TTATCCAGGG ATCCACCCCA TAATCAGCAT CCAAACGCTG 6360
 ACACCATTGC ATACGCTAGC AAGATTTTAT CGAAAGGACC CAGATGTAGC TGTCTCTTGT 6420
 GAGACTATGC CGGGGCCTAG CAAACACAGA AGTGGATGCC CACAGTCAGC TAATGGATGG 6480
 ATCACAGGGC TCCCAATGGA GGAGCTAGAG AAAGTACCCA AGGAGCTAAA GGGATCTGCA 6540
 ACCCTATAGG TGGATCAACA TTATGAACTA ACCAGTACCC CGGAGCTCTT GACTCTAGCT 6600
 GCATATGTAT CAAAAGATGG CCTAGTCGGC CATCACTGGA AAGAGAGGCC CATTGGACAC 6660
 ACAAACCTTA TATGCCCCAG AACAGGGGAA CGCCAGGGCC AAAAAGGGGG AGTGGGCGGG 6720
 TAGGGGAGTG GGGGTGGGTG GGTATGGGGG ACTTTTGGTA TAGCATTGGA AATGTAAATG 6780
 AGCTAAATAC CTAATAAAAA ATGGAAAGGA AAAAAAAAAA AGAAAAGAAA GAAGCTACGT 6840
 CTCTAGAGAA AACTTTTTTT TTTTTTTTTT TTTTTTTTG GTTTTCAAA ACAGGGTTTC 6900

TCTGTGTATA GTCCTGGCTG TCCTGGAAC TACTCTGTAG ACCAGGCCCG CCTATGCCTC 6960
 CCAACTGCTG GGATTAAAGG CATGCGTCAC CACTGCCCCG CCAGGGGAAA CTTTGAGACC 7020
 ACAAGAATGA AGAGGTCAGA GCCATTTTCC TTATGAAGGA GGCTGAGGCT CCATTGAGGA 7080
 ATTGTGGGTA TGCTCGGATC TCAAGCCTGG TCACTTGGAT GGCTTCTTGT AGAGACCTTT 7140
 AGCTGCATCT GTCTCCAAAC TGCTTCCCAA CCCCTGGAAC GGGCTCTGAA GCTGTCCCTG 7200
 CCTATAGCAT GCAAGGCCTT GTGAGTACCA GGTATGAGGC CTGATTGCTA GAGAAGACAG 7260
 GATCTCATAG AGTCTCTTGC TATTTGCAAT AGGGATCATT CTTGGAATAA TCCGAAAAGT 7320
 AGAGTTTAAG AAATTTTGAA GAAAAAAAAA TCTAATATTA CAGATTCCAG ACTTGTTATA 7380
 TAGAAGAAGA AGAAGAGGAG GAGGAGGAGG AGGAGGAGGA AGAAGAGGAA GAAGAAGAGG 7440
 AAGAAGAAGA GGAAGAAGAA GAGGAAGAAG AAGAGGAAGA AGAAGAAGAA GAAGAAGAAG 7500
 AAGAAGAAGA AGAAGAAGAA GAAGAAGAAG AAGAAGAAGA AGAAGAAGAA GAAGAAGAAG 7560
 AAGAAGACGA GGAGGAGGAG GAGGAGGAGG GGGGGGGGAA GAGGAAGAAA GAAGAAGAAG 7620
 GAGACGGAGA GAAGAAGAAG GAGAAGGAAA AAGAGAAGAA GAAGAAGGAG AAGGAGAAAAG 7680
 AGAAGGAGAA GAAGGAGGAG GAGGAGAAGG AGAAGAAGAA GAAGAAGAAG AAGAAGAAGA 7740
 AGAAGAAGAA GAAGAGGAGG AGGAGGAGGA GGAGGAGGAG GAGGAGGAGG AGGAGGAAGA 7800
 AAAGTGAACA GTAGGGATTG GAGAGATGGT TCAGTGGTTA AGAGCACTGA CTGCTCTTCT 7860
 GGAGGTCCTG AGTTGAATTC CCAGCAACCA CATGATAGCT CACAACCACT TGTAATGGGA 7920
 TCCGATGCCC TCTTCTGGTG TGTCTGAAGA CAGCTATAGT GTACTTGTAT TAATAAAAAT 7980
 AAATAAATCT TTTTAAAT TTTTTTTAAA ATAATGTGAA CAGTAACTGC TGTCTCCAA 8040
 GTGCCCCTGT TGTCATTTTT AAAAAGCCAT AGTTCTTTCT TTCATGGAGG GTGATCAATC 8100
 ACAAGGTGCA CTGCATACAT CTAGGATAGA AGCTGTGTTA CATAGATTCTG GTGTGTGGAG 8160
 AGTTGCTGAG TCCTCTCTT TCCTTCTTTC TCAAAGGTAT CAGCCAGGCG TCATAGTCCC 8220
 ATCTCGTGTC TCAGGCAGCT ATCCTATCTT CTCTCCCTC TTTGTGACAT TGATGACCAT 8280
 TCATCCAAAC AAATGGAAAC ACTTCCCATG GGCCATTGAG TGCAAGTCTT CCACGTGGCC 8340
 TTGCTTTGTG CTGGGGAAGA GTGTAGACCT CAGCTGTCTC TTGAATTCTG CTAGGGCCTG 8400
 GTAGTCTAAA CTGCCAGAAG GCAGCAACCT CTGCATTTTG TTCATCCATG TGGCACCAGT 8460
 CAGTGTGAG AGAGAGAGAG AGGAGAGAGA GAGAGATTAA GTACAGTCTG TCTTTGCAGA 8520
 TCCTTGAAGA GTGGTTTGGC CGCACTATCA TCCGTTCTTG CACCAAACCTG AGCTACGACC 8580
 ATGCCCAGAG CATGATCGAA AATCCAACCTG AGAAGATCCC TGAGGAAGAG CTCCCCCAA 8640

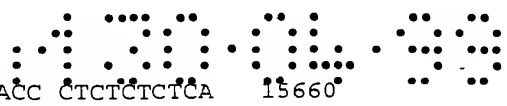
TTTCTCCAGA GCACAGCGTC GAGGAGGTGC ACCAGGCAGT CCTGAAGCTG CAGAGCATTG 8700
 CAAAGCAACT CCGCCGCCAG CGCTTTGTAG ATGGCGCACT CCGTTTAGAT CAGGTCAGTG 8760
 AGTCTCTTTT GTTTTATGTG GTCTTGAGTT TGGCTTGTGC CCAAACTCA AGGGTGAGAA 8820
 ATATCCTGGT GGCCTCTTTC TCTCCACCTA TTTCCCCTGC CCCTGCCACA CCATGGTAAT 8880
 ATGAGTTAGG GTAAGATGGT ATCTGTGTAC AGAGTTCTGT GACTCCCAGC TGCTCTTACC 8940
 TGGAAAACCT GTGTCCATGA TTGAATTCTC ACTTGTAGAT GGCATTGCTG TGACAGGTCC 9000
 CTGGGACAAA GAAGGGAGGA AGGACATATT TTTGGCTTGT GGTTCAGAG GCTCTTGAA 9060
 CATAGCTCTG TTGTTTCTGG CCCATAGTTG GGGGCGGGG GTGGCATGTG AGAAGTATGT 9120
 GGCCAGTG AGCTGCTTGT CTCATGGCAG CCAGTAAGCA GAGAGACAGA GGCATGTGAA 9180
 GGAGCAGAGG CAAGATAGAC TTTCCAGGGT ACACCCCCAG TGATATCAAT GAATCCAACA 9240
 GCTGGTTCTT TGAGAAGATA AGCAAGATTG ACAGACCCTT GGTCCAAGTA GCCAAAAGAA 9300
 ATAAAGAAGG CCCACATTAA CAGAGTCAGA AATGAACAGG GAAACATTAC AACAGATGCC 9360
 TAAGAAATTC AGAGTTTCAT AAGGGCATACT TTTAAAAAAC TGTACTCTAT TAGAAATGGA 9420
 TGAGTTTCTA GATTTCAGCCA AACCACCAA ATTAACCAA AAAGAAGTCA ACAACCTAAA 9480
 CAGACCCATA ACAAATAAGA TTGAAACAGT AAAAACAAAA CAAAACAACA AAAAATTCC 9540
 AGCTACAAAG AAAAATCTAG GGCCAGATGG ATTCACAGGA AAATTTTACC AGATGTTCAA 9600
 AGAAGATTG CACCGAGTTG TCCTTAACT ATTCAAAAAG TAGAGGCAGA GGGAGCACTC 9660
 CCAGGTCTCC TCTGTGAAGC CTTTATGTCA CCAGTTCTCT CCGCTCATGG AGATTACTTC 9720
 CTCTGCTCCT TGCTTCATGC TTGGTGTCTT GAGGCTGCAG CCCACCATCC TGTCATCTCC 9780
 ACCAACAGTC CCTCCCTGAT TCCAAGAGGC TAAGTTGATG CTAATGACAC CAGAACTTGT 9840
 GTCTGACCTT TCTCCCTCAC TCAAGCCTAG CTTCTTTACC TGCCTTATCT GCCTGACTGC 9900
 CCTTCAGCAG CACAGTGGTG CTCACTCACC CTTCTTCTG CAGAAAGCAG TGCTTGATGC 9960
 CCACAGCATG GCACACAGGC TTCCAGCAT CCTCTTCTCC CACTGATACA CTGGAGCATT 10020
 ATATATGTGC CCCCAACCCA AGTGTAACCAG TCGCACAGAT TTTTGTAATT ATGCTTAGAC 10080
 TAAACATTAG ACAGACAGAT CATATACAAC TCTCAAAAGG AAGCTGTTTA TTCTGTAAAC 10140
 ACATCCATGT TTTAGAAAGA CAAGTCTTCA GAATGTCTTT AGGAAGACTG AAGTCACTTT 10200
 ACAAATGAAC CGTGGGGCTT AGGAAAGTCT TTAGAAAATG AATTGGGTTT AGTTTCTCA 10260
 AAAAGACTAG GAATCTATGA TGTGGCACC TATAATCTCA TCTCTCAGGA AGCCAAAACA 10320
 GGAAGATTGA AAGTTCAAGG CCATATAAGA TGTATGTCAA GATCATGTGG CAAGGAAGAA 10380

TAAGAGGAGG AAGCAGAGGA GGAGGAAGAG GAAGAGGAGG AAGAGGAGGA AAGAGGAGG 10440
 GAAGAGGAAG GAGGAGGAGG AAGGAAGGTG GAGAGAAAGG CAATAAAAAG AATAAATTTA 10500
 GTTTTCTCTC ACTCTGTAGC TCAGGTTGAA CTTGAACTCA TGGCTAGCCC CCTGCCTCAG 10560
 CTTCCCAAAT GGTAGGATTA TAGGTGTGAG CCACCAAACC AGATACTAAC TTGTATTCTT 10620
 TAAGTCTTAC TTTTTTTCAA AAATGGTTTA GAAACATATA TCTATGTAAA TTAAGTTATA 10680
 ATACAAAATG TTAGGTTGTA TATTATGTAT GCCTTTTCTG CATGATTCTC TTATTTACTT 10740
 AACTTTTACA ATGAAAAACC AGCTGTTACC CAAGCCCATC AAATGAGGAA GTTTCTGAAG 10800
 TACCATTTC AGATGTTTC CCACTAAGAT GCTATAATAA AATTCAACTG GATTAATTCA 10860
 TCTGTGAAAC TGGAGGGAGG GGGAGAAAAT AGCGGCAACT TATCTCTGTC CCATTGGAAG 10920
 AGGTGTGGTC ATCATCGTAA TGACCATAGA TTATTGATGG AGAATGAGCA GTTAGTATGT 10980
 CTGATACTCA GAATTGTATT ACTGAAAAGA CTTTAGATAT CTGTATCCCA GTGGGCCTCC 11040
 TAACTCATAA ATGAGAAGGC TGAGGTCCCC ACAGGTAGAT GGGTTGCTTA TTGCCAGGCA 11100
 TCCAAGTAGC TCTTTGTTTG GTTTTCCTCC ATTTATTACA CTATGCTGAC ATAAGAGAAA 11160
 AAAGTTTGCC TTTAAAGTGA AAGGGGAAAA CACCCTCAAA AACCTAATTA GGTTCAGTT 11220
 AATTAAGGTT TGAAAGTAAT GAATTGTAT CCTTGGAGTT GATCCCTTCA TTCGCCAGAA 11280
 AACAACTCTG TAGACCCCCA CATAAGATGG AGACATCAAT CTTTGCAGCC AAGGACACTG 11340
 GTGAGGCCGT TTATAAATCA GCTAAATGGC TTTATTGAGA AGCCCTGCGT TTGTTCTCCC 11400
 GTCCCTGTTG CCTTCTTTGC CCTCACAAGT TCATTTTTCC TTGGTGCCTT TTCAGTGGCC 11460
 TGCTGTTTGC CATTGTTCTC TGAAGCTTTG TCTGCCATAG TTTACTGTGT CCATGTTTTG 11520
 GGTGGTAGTC CTTTAAAAAG CACATCCTTT TATGTCAGCA GCAATTAGAG ATCGGTCTTC 11580
 AGCCAATCCA AAGGCTTTGC CTTTCAAAAA AACAAGGGTT GAAGAACCCG AAAAAGAACA 11640
 AAGAAGAAAG CCAAGCAAC AAAAAGGGGC CTGGTTGCAA AAGCAAAAAA AAAAGCCCCA 11700
 AAAAGGCAAA AAGGCAAAAC AAAGTCCCA ACCAAAACCG AATTAAAAA AAGTTTCCTC 11760
 CAAAAGGTGA TTCTCCTTTG CCCCAAAAGC AACACAGGCT TCCAAGGCTA TCTAGTGATT 11820
 TTTGGTCGCT GAGTTGAATG ATGACCCTTC TGAGTGGCTT GTCTCTGAAT CCATGTTTTC 11880
 AGCTACCAGG GTAGTTCAAG GACTTGGTAC AAATGACCAC TTTAATTATT TGTTTATAAT 11940
 ATATGTCTCT CCCGAATCTT AAAAGAGGCC ATAATGGGGC CAAGACTTCT GTATCTGTAG 12000
 AAGAAAAGGA ATCAGAGTGG TTCCTAATAT CCATATACTG AGTTTGATGC AAGGGGAGCC 12060
 ATCTGAGGGT TTTTGCTCCT GACTAGCACA GGCCAGCCCT CAGCAGCTGC CATCTAGGGG 12120

GGAAGATAGA	TCTGCCTGGC	ATGGGTGTAT	TTAAAACCCT	GAAACCCCTT	TGGGCTTCTA	12180
GGTCAGCTAT	TGCCTTCAGA	AAGGATATGA	TGGTAAGGTA	ATGGGGTGCC	AAACAGATCC	12240
TCAATATAAG	ACTAACATTG	GCTGATGTCA	GGAAACTCCA	CGCCCTGCTT	TCTGAAGCTC	12300
TCTGAACCTG	TTTCTCTTCA	GCCAGGCTAA	GACTTCTATG	TGAAACAAAC	TAGAAGTTTG	12360
CAGAGATCAG	ACAAGTTCTC	CCAGCAGGCA	GTTAAAACCTA	TGAATTCGGA	GGGCCTTGGA	12420
AGTCAAATGA	AAAAAACCTG	AGAAAAATTC	ATATAAAGTA	AAGGAGGCTT	TACTAAGTTC	12480
TCAGCTCTGT	CATCTCTGAA	ACCTACTTGA	CACAGTTTTG	AGGCCCAAGC	TCCATGCAGT	12540
TTCTTTGTAA	AGGTAGCCTT	TCTAATGGAA	GACACTTTTG	AATACCCTGG	GACTCAAGCT	12600
GTGTGAGTCT	GTAATGTTTG	ATCCTAACCT	AGCATAGCCT	TTCAATCAGT	GTTGGCAGGC	12660
TTTCCCAGGA	AAGGCCAGAC	AGTAAATGAC	ATGAGCTCCT	GGTCCATATG	GTCTGTCTCT	12720
GACTCAGCCC	TGCCTGTAA	TGTGCTCCAA	ATGAATGGGG	GTAGTTGAAG	GTCACTAAGA	12780
CTTGGAATTTG	ATATCATTTT	CACAGACCAC	AAAATATTAT	TCTTCATTTG	ATTATTTTTTC	12840
AAGTATTTAA	AAATGTAAAA	ATTCTTCTTT	GCTCCCCGGC	CATGCAAAGC	AAGTTAAACT	12900
GTGTCCCACA	CATCACTGAC	CCTGCTTAAC	TGACCAACAA	GCTTTTCAGC	CCTATTACCC	12960
GCCAAGCCCTT	GAGCAGCTCA	TTACCACTTC	CCCAGGAAGC	CAGGCTAGGA	AATGGAGAAC	13020
AGTTGGGCTA	AGTGACTTCT	CAGGATGGTT	CCATACAATT	AAGTAAATTA	TTCTTTTGAT	13080
TAGTACCACG	CTTAGGGGGC	CAGTTGGAGG	CTGGAAGTAA	GAGTGACTGA	CCCCCAACC	13140
CCAGCACAGT	TCTTTTGCCC	TTCCCAAGGT	CCAGTCCCTT	TAGCTTGAAG	CCAAAGAGTC	13200
AGCACTCTCT	TTACTCCTCT	GCAGGACCCT	CAGGGTCAGA	GCAGCCCTCC	CTCTCCCCTC	13260
CCCTAGCTCC	CCCTTCTCCT	TCCCTCCCCT	GGTCCTCTGA	AGGTAGAGAC	TACTCCAGGA	13320
AGAGCAGGCT	ATGAGGAAGG	TGGGTAGCTT	CTCTCCTGGC	TACCTGTCTG	CAGTGCTAAT	13380
TACAGCAGAG	TGTTCCCTCT	CTCTGCCATA	GATAGCTGCA	TTCTGGATGG	CTGCTGCTCA	13440
GTGTTGCTCT	CCGATGACAT	TGGTGTAGCT	GTGGAGAATG	GGCAAGCCCT	TCTGGTTTCC	13500
TTTAGCTTTA	GTGTCTGTGT	CAACTCAAAG	TACAACATAG	TCCAAGGCCC	AGGCTCTGAG	13560
GTTTTTCATT	CAGAGAGTTC	TTCACTCAGC	ATAGCTTCAG	AGACCTGTTT	GGGGAGCCCA	13620
GTGTGTGTGG	AGGGGGTGAG	AATGTAAATG	AGGAATGAGA	AGTTTCAGGT	ATGGGAAGGG	13680
AGGCAGTGAA	CCACTAGACA	GTAAGAAGCA	CTGGGTGGAA	GTGCTTGCTG	AACTTGAAAC	13740
TGAGGAATGA	CTCCTGCCCC	AAACCAGTGC	TCATCCTTAG	AACCCTGAAG	AAATCCATGT	13800
GCCTGAAGCA	TACTGTCTTA	GTTAGGGTTT	TACTGCTGTG	AACAGACACC	ATGACCAAGG	13860



CAAGTCTTAT AAAAAACAAC ATTTAATTGG GGCTGGCTTA CAGGTCAGAG GGTTCAGTCC 13920
ATTATCATCA AGGTGGGAGC ATGGCAGTAT CCAGGCAGGC ATGGCCCAGG AGGCACTGAG 13980
AGTTCTATGT CTTTCATCCAA AGGCTGCTAG TGGAAACTG ACTTCCAGGC AACTAGGGTG 14040
AGGATCTTAT ACTCACACCC ACAGTGACAC ACCCATTCCA ACCAGGTCAT ACCTATTCCA 14100
ACAAGGCCAC ACCTTCAGAT GGTGCCACTC CCTGGTCCAA GGATATACAA ACCATCACAC 14160
ATACCAAGAG CTTTCTGTCC TCTCTGATCT TCAGAGGACA TCATTTGTAA CTCCTGTCTC 14220
TTTGTGCCTT TCACTTCCTG TAATATGTCA CAGGAGTCAT TTGTGTTGAC CGAAAATCCC 14280
TCTGTTATTT ATCATAACACA CACACACACA CACACACACA CACACACACA CACACACACA 14340
CACACACACA TACACACACA CACAGTAGCT CTGCGACTCT TTAGGGTAGT GACAGTGGTT 14400
CAGTGGGCTT CTGCTACTTC CAGGCCTTCC ATTTAAATGT AGACAGCACA TGGCTTCACT 14460
TGGATATTTA GCAACTCACT TATTTCTCTA CTTTCCTGCT TATTTTCATT TGTAGATCCA 14520
GCTCTCTGTG ACACTCAGAC CTGGACTCTC AGGGGTAGCA GGAAGGGTGG GGAGCTGCAC 14580
CCTTCACCAC AGAGAATCAG AACACAGCCT ACAGTGGGGT CTGGAAACCT TTCCTTTGAG 14640
AGTGACAGAT CAGTTTAGTT ACTGTACATT AATTTTCATAT GGAATTACAG AAAATAGTCA 14700
TACTTATGCA CACATCCTTC CTTGTTAGAT GAATTTCTCT GGGTGGCTTG TTAGTACCAT 14760
CTGCGCTCTC CCTATACTCA CTCTCCCTGT GACACAACAT AGAGCCATTT CTCCCCTTC 14820
CAAAAACCTC AGAAAATCCT GTTTACCTTG GAAGTTGTTA TGAATGCAGA CTGACACTTG 14880
ACCAGTGGCC ATTGCTAGGT GCCTCTTGAG TTCTCTCTCC AACAGCAGGA ACACTGCTCC 14940
TAACACTGCT CCTACAGCAG TGGGAAGCAG ATGTCCTACC CTAAGACTGC ATACCAAGTA 15000
GAGGAGAACA TATGGACTTA GCAAAGGAGG CCGAGGGGAT CTCAAGCACG ATGGGGAGTG 15060
GATGGGAGTG AAGGGCAAGG ACAACCTGCT CAAGACAGCT GTGCCCCACTG ATGAGCATGA 15120
GAAGAGCCAG AGGCAGCTTC TCCTCCTCTG AGCTGAGGCT GAGACTGGAC ACTTGTGACA 15180
CACGGAGGTG AAAGTGGCTC TGTCTACCCC GAGATGGTTT AGATGAAAGG AGGCAAAAAA 15240
GTAGCCAGAG ATAGAGCCAC ACCCTCTGCC AGCTGGAACA CTTGGGATGC TTCCCCTCTC 15300
CTCCACCTCT GCTATTACCT TGAATGTTGG GTGTCTTTCC AGGCAGGATG TAGTGAGGCC 15360
TGAAGCTGGA ACTGCTGCAG TTGGTCAACA GGCCTGTTCA GAAGAACACT GAGTCTGCTT 15420
TCTAAGTAAC TCTAGAAAGC AAGTTTGGCT CCTAGCCAC CTCTAGAAGC TTTTGCTTGC 15480
CTTCTGGTTC ACTCTGCATG TTGATGTCTA GCCTCATTTT TTCCAGGCCA AAAAAAAAAG 15540
CATTGCTTCA TGCCTGCTGC TATATTCTCT GGGTTCACCT CTCTCTGGAC CTGAAGAATC 15600



TGAATACTGA	AATCCTCTGC	TTGTTCCAAG	TGGGGCTGGC	TCGGCCAACC	CTCTCTCTCA	15660
GGGTGCCATA	GCCCTTCATG	CCTATCTTTG	TCACACTGTC	CAGTTGTCTT	GTTACCCCCT	15720
CTCTACCCCCT	GTCTCCTCCC	CTAAGATTCA	GTTCTACAG	AGCAAAGACC	ACATGCTATT	15780
GATCTTTCTA	TCCTCACTTC	CTGAACAGTG	CTGCATTTTA	ACAAGCTGTT	TGTTCAGGGT	15840
CTCTAAACAG	TGCCATGCAT	GCTGGTCTTT	TTAAATAAGG	TACTGCTAGC	TACAGTGGGG	15900
AGAATGGAAA	CCAAGGCTGT	AGATCAGAAT	GTTTGCATGA	GAGAGTTACT	ATACAGTGTG	15960
AACCAAGGCT	GCCCAAGTAA	ACTGGCTGTT	ACTTAATTCT	TTGCCAGGGC	ATCCAGCATG	16020
TAGAAGAGAT	GTGGTGAGGA	CTTTCTCAGG	TGGAGCTGTC	CTGATAGGCA	TGAGGAGTCA	16080
GAAGGCTTCA	GTATGCTTGG	GGTCATCGAC	ACTTCAGAGG	TTCCCCCTCA	GATTGGGATG	16140
TCCCTGCTGG	GGATGTCAGG	AAGGACACTC	CCAAAGTTCC	ACCAGAGAAG	AGAGATGCTG	16200
GTCTAAAAAG	GCAAAAATTA	CCTCCTCCCA	GAGCTACTCC	TCTTACCTCT	GGAATGGGGC	16260
AGAAACAAGT	TGGATAGGAA	TGGCAACCTC	TAGTCTTTGC	AGGATCCTGA	GAGGACTCCA	16320
CCCCTACCCC	CACCTCCGTT	TTGCTCAGAA	TGGAAATGGC	GGCTACCAGA	TAAAGACTTT	16380
CTATTGGTCT	TTGGGGCTTT	TTAAGAAGAG	AACTTAAATA	CAACCCAGGT	TACTCAAACA	16440
GAAGTTGCTG	ACCTTCCCAG	GGTACAGTGG	AGGGGAGGAA	GGGCTCTCAT	GCTGACCAGA	16500
AGAGACAAGA	ACTTCTGTGA	CTTAAACAGG	GCATGGCTAG	AACCCTCATT	TCCTCAGAGA	16560
TGAGATTATT	TTGTCTTATG	ACCTTGACAG	ATGGAATGGA	ATTTGGCCCT	TCTGGGACTT	16620
TGCCTTTTGG	GTAATTGTAC	TCAGTTAGGC	AACCCTGGGA	CTCTCTTTAT	TCATAGGACA	16680
TACTGCATAT	TCTTGCCCTG	CCCCCATGTC	AACTCACGT	CAATTGAATG	TAAGCCAGAC	16740
AGCTACATAA	GAAGCATGGA	ATGCTTTGAC	GTTGGTAAAA	CCTGCATTGG	AGAAAGAGAA	16800
CCCTTGACAG	TGATCCTTAG	ATTTCAACCA	TGACTGCTTC	TTGGGACTGG	CCCAGTTGAT	16860
TTCAGTTTGT	ATTCTTCAGT	GCGCTCGGGA	CTCTGTTTCC	TAGGCCAAAG	CTCTTCTGTT	16920
CTGTTCATTC	TACACTGAGC	TCCTGCAAAT	GTTCCCTTGT	CCCTCAAGAA	CCTGCGGGTA	16980
TCACAGACCA	ATGGCAGAAA	TGTCTGGGGG	ACAACATACA	GGTGTTTTAT	TTACCACAC	17040
AAGGATATAT	TAAAAAAAAA	AGTTAGGGTA	GTGGTGCCCC	ACGCCTTTAA	TTCCAGCACT	17100
TGGGAGGCAG	AGGCAGGTGG	ATTTCTCAGT	TTGAGGCCAG	CCTGGTCTAC	AGAGTGAGTT	17160
CCAGGACAGC	CCAGGTTATA	AAGAGAAACC	TTGTCTCAGA	AAAAAAAAAA	TTACTAAGCT	17220
AGGGCTATAT	AGCTTAGCTG	TTAAGTGCTT	ACCCAACAAC	ATGAGACCTT	GGGTTCAATT	17280
TGCTGCACAA	CATAAACTGT	GTAGTGGCCA	CACACCTGAA	ATCCCAGCAC	TCATGAAGTA	17340



GAATCAGGAG	AATCAGAAGT	TCAAAGCCAG	TTTCAAATAC	AGAGAATCTG	AGTCCAGCCT	17400
GGAGTGCATA	AAACCCTGTC	TGGGAAAGAA	AAAAAAAAAA	AAAAAAAAAGC	AGTGTTCCCG	17460
TACACATGAA	GCATTCTATC	CCCAAGACAA	AGGAAATACA	CGATGTGACA	ATATGAAGTA	17520
GGTTTCTAAT	ACATTTTTAG	TTATTTGGGG	AGTGTGAAGA	TATGCATCAC	AGCACACAAA	17580
TGACGATCAT	AGGACAGCTT	ACAGCAGTCA	GCTTTCTTCT	TATACCACAT	GGGTCCGAAG	17640
ATGGAAGTCC	AGTTGTGAGA	CTTGGCCGCA	GGCGAGTTTA	TCCACTGAGC	CTCTCTCCGG	17700
CCATGAAGCA	GTTACTTTAC	GTTGACTCGC	TTGAGCTTGT	TGGGAGCATG	CTTAATTATT	17760
GCTTTGCTCA	CTTTGGTTGC	CTCAGAGTAG	CTTGCGAGAA	TTACTAGACT	CACACGTTAG	17820
ACCCAGATGT	CTTCTGCCTT	CTGATGAGGA	GCAAGCGTGT	GAGTAAGGAG	GGGAAGCAGG	17880
TCACAGTCCA	AGCCGCTCAA	GTCTGAGCTG	CAAATCCTTC	ATTGTACAGA	CGGCTCCGAA	17940
TCAGAACT	TCCTGTTGCT	ACAGTCAGGA	CGGTTATAGT	TTTTATTGTT	ATAAATGACA	18000
TTGTAATTAA	TACCCTTACA	CAGAAAGTGT	AAAAGTCACT	TAGAAATACA	AACATCATAA	18060
ACTACTAGGT	TGAAGAAAAT	TGACTTTTTTC	TGTGTCAATT	CTTAAGATTA	ACTTTGATTA	18120
TTTTATTGTA	AAATGAATAT	ATGTTCATAC	TGTAAACATA	TTTAAATAAA	CAAGGAAAAA	18180
GTAGCCATTG	GCTATGCCTC	ACCTAGTAAT	AATACTTAAT	ACTGTTCACT	TCAGAGCTTT	18240
TGGCTTTCTG	GGTGTTTTCC	AGAAGGTTGG	ACTAATTGAG	GTTTACCCCA	TCAGAGAACA	18300
GTGCTATGCT	GTTACTCTTC	TCAGCAAATT	CAGTTTGTGG	CTTTGCTTTA	ATCTTTGTTA	18360
GTGTAAGTAA	CTTGGAAGTG	GTGTTCCATT	GTTTGAGTTG	CCTTTTTTCC	TCCTGTGTCT	18420
CTATCAACTC	TCAGGCCTGT	CTTTGCCAGG	TCTGTGGAAA	GCAGATGCTA	CATCCCATCC	18480
CTAGGACTGC	CAACAGCATC	AGCACAGGCC	CCTGCTCTGA	TCAAATACAA	CCACCTTTTT	18540
CCCTATGAAG	ATAGAATTAT	ATACAATAAA	GTCCACCATC	TTTAGTGTAT	AGGTCCACAA	18600
GCTCCACACA	TAATCATATG	TCTACCATGG	TCAAATACA	GAATAGTTGC	CTCACCCAAT	18660
AAGCTCCACA	TGTGCCCTTC	GGTAGGCAGA	CTGTCTCACT	TATCCTCAGT	CCCTAGTAAG	18720
CCACACATGA	GCACATGCAT	ACAGGGTACA	AAGGTCAATT	TAAGGTACCA	TTCTTCAGGT	18780
GCCCTCTACC	TTGTTTGTTG	AAACCGGATC	TTTTACTGAG	ACCCAGAGTC	ACCAATTGGC	18840
TCGCCTATCT	AACAGTAAGC	TCCAAGTATC	GTCCTGTCTC	CTCCTCCCCA	GCACTGGGAT	18900
TACAAGCATG	TGCCACCATG	CCTGGCTTTT	AATGTGGGTT	CTGGAGACCA	AACCTTAGATC	18960
CTCATGCTTG	CATGGAAACA	TGTTCCAAC	GAGCTATCTC	CCTATTCTAA	TTTTTGCCCCA	19020
TTTCTTAGGT	GGGTCTTTTG	GTTTCCTAGT	ACTAAGTTTT	GAGGATTCTT	TTGCTATTTT	19080

AAATAGAACC	TCTACCAAGT	TGTGTGATAC	TACAAGCCAT	CCAGCTCATT	CTTTCATCCC	19140
TTGTCTTATT	CTTTCTGGCT	CTTCTTTATT	CCCTTTCTTT	TGAAAAGAAG	TTTTTAATTT	19200
TGAAGCAGTC	CAGTTTACCA	ATTGTGTCCT	TATGTTATCA	AATCTAAGAT	TTTTGTTTTG	19260
TTCGTTTTGA	TGGTATTATT	ATTTATTATT	ATTATTATTT	AATGTATGTG	AGTGCTCTAT	19320
CTGCATGTAT	ACCTGCATGC	CAGAAGAAGG	CATCAGAACT	CATCATAGAT	GGTTGTAAGC	19380
CACCACGTGA	TTGCTGGGAA	TTGAATGAGG	GACCACTAGA	AGAGCAGACA	GTGCTCTTAA	19440
CTGCTGAGCC	ATCTCTCTAG	TCCTATTCAT	TTTTTTTTTAA	ACAGTCTTGC	TATGTAGCTC	19500
AGACTGGCCC	CAAACCTCAAG	ATCCTCCTGA	CTCAGCTTCC	CAAGTGCTGA	GATTACAGGC	19560
TTGTTCCCTCT	AACTCCTGGC	ATGAGAAATC	TTTAACTGAC	CTAGAATCAC	AGATTTTCTT	19620
CTAGAAGTCT	TATAGCTTCA	GAATTTATTT	CTACTTTCTC	TCTTCCTTTA	TAAACACATT	19680
CCTAGGCCCA	GACATTTCTT	TTGGAAAAAA	GTTCCAATAA	CAGAACTGGA	CACACCTGAG	19740
CAGATGTAGG	GTAGAGTCAG	ACCTGGGAGT	CTTGCCAGGC	ACAGTACCCT	CCTGGAGCCA	19800
TCTGCAAAGA	AGTTACCTCA	GGAGTGGCTT	GTAAGCAGAT	CTTCTCTGGT	TTTAAAGACT	19860
TGGCATAAAA	CTGAAAAGTG	TATCTTTTGA	ATCAGGGAGC	AGAACGATAA	GAGAGAAATC	19920
TCTCAGCTCT	CTAGACAAAT	CCTCTTGACT	ATCACAGAGC	TGATGGTGAG	CGGAGCCAAG	19980
CAAGACTTTG	TCGATTACAT	GCAAACGCCC	AAGTCAGTGA	CTCACTCAAT	CATGCTTTAA	20040
TCTCATAACT	CAGTGGCTTT	AAAAATTACA	GTCAACAAGG	CAGCTCGTGG	GTTACAAGTG	20100
CCATTGGAAC	TAGGTTTTCT	CTGAACAGCT	GGAGTGTAAT	GTGGTGGGAA	GAAAGCCTGC	20160
TGTGGGTGAG	AGGCCAAAGA	CTGTTTGCCT	GGGAAGGATG	TGCAACTAAC	GTTTGATAAA	20220
AATCTGTGAA	ATGACCACCC	TCAGCCAATC	TAAGTAGAGG	CCTGCCATTT	TCATCCATGG	20280
GAAAGTGCAT	CACAGCAAAA	GCATTCAGAA	GGCACTGGTA	AGACAGTGGC	AGTCACCATT	20340
CATCAGACAA	GACAGCCCTG	ACTTCAGGAA	GTGTCAGGAG	TCAGAGTATG	AGTATGGAAT	20400
ATTAACAGAG	CAGGCAGAAG	ATTCCAATTC	TAGTCAAGGA	GGGCCAGTGA	GAGAGAACAG	20460
TTTGGAATG	GCTTCTCTGA	ACAGATCCAG	GCAGATCAGT	GCAGTCATTT	GCTATGTTCT	20520
AAAATGTGTA	GGCCTCTGCC	ATAGCTGTGT	CACGGAGGAT	ATATAAACAG	GCTGTTCTTT	20580
GAGGACCTCA	TTGGGCTGTC	CCCAGGCACA	AACATTTTCT	TAATTTCAAT	GTAGAAGCTG	20640
TTACCCACAG	GAGAGATGGA	GTAGGACTTT	GGTTTCAGAG	CCCTATCTAT	AGCAGCTTTG	20700
TTGAGACCTA	ACTGGAAAGG	CTCAAGATAG	GACATCACAC	AAGGCATTTA	GAAGCTTGTA	20760
GCAGTCATCA	GACATCAGAC	CAGACCTGAC	AGGAAGAAAC	AGGTGAGTCT	CAAGAGGGTT	20820

CATCAGGATG CTCACGAGTT TCTGCCTGCA CAGCATGGGC ATATGGTATT ACCAGGAGAA 20880
 GCCATCTATC TGCCCATAGG GGACAAGCAG ACATCAGTTG GGTGATAGGG ACATGAAAAC 20940
 TTTCTGGCCC ATCTTTATAT CTGTTCCAGT GAAAGATGTG TGAGGTCCTC ACCCCTGAAG 21000
 GCTCTATACT TCCCTCTCCT GCTAGACAGT CTAGCGAGAC TAGGAAGCAA CACAGAATCT 21060
 AGATGAGGCC TCTGTGAGCT GCCCAGGTCC TTAGGAGTGG AGTGGGGCAG GACCCGTTAC 21120
 AAGAGTACAC CCCCCGCCCC CGCAATGAGC CCAGTTGTTC ACTATGGGGC CGGGAACATC 21180
 ACCCAGCAGG CCCTATTGGT CCTGGCCTGC TCCCCTCCCT CCTTACCTCC TCACTCACTC 21240
 TTCCCAGCTC GATCTTTCTC GCTTGTTAGA GAGAGAAAAA AAGTGAATTC ACTCCCAGTC 21300
 CTTTTGAAAC CCAATGTGTC AGTGATCGAT GAGGCTGTAT TCTCTAACTT CAAAGGAGAA 21360
 AACTAAGTA GAGTGAATAC TGGCCAGGGG AGTTGAAAAG TCCCAGGGAG TAGGAGACAC 21420
 AGGAGTGACC CTGCCATCAT GAGGAGCACC CCCCATCCCA CCCCTGCTGG TGCCATGCAG 21480
 AAGCACAGAC AATGCCACTT TCAGTAAATC ATGACGGATC CTGAATGCCC AGTTTTGTCC 21540
 TGTTTTCAAT GGGCTGTGGG CATATTGCTT AAGATATAGC AAGCCATTTG TGCTGGGTTT 21600
 CCAGCTACTC AAAGGCTCGA CATTTGAGTG TTCTCTCAAT TGTATAATAG AGCCTTTGCA 21660
 TATGTGATTT GGGGGGAGGG TTTTTCCTC CAGATTTCCTA TAGCTAATCA TAGTAGAGGT 21720
 GACCTCAAGT GTAGTGCAGA CCATTGTCCC TCTTCACCCC TGCAGATCTT AGCAGTGCTG 21780
 AGCTTTAGGG ATATTCAGGC AGCACCTAAT TCAATCACAC ATCTGACCCC TGCCTCTTTG 21840
 GCCACTCCTC TGAAACTCAG TTAGCTCCCT GGGGTCTCCC ACCCCACAAG CCTGGATCCT 21900
 CAAGAGCCTT TGTACTGAGT AGAAAGTGCT CAGACCTTCC TCCACCCTAT CCAGATTCCC 21960
 ACTCCCCCG CCTGAATTTA AGCACAGAGA ATCCAGTGCT GCAGGGCCAC TTGTTCTCAC 22020
 AAGGCTGCAC TTGTGGAGAT GCCTGTGTGA AGCACCTGT AGACATCCCA TGCTAAAGTC 22080
 TTGGGAACAC AGAGAAAGAA AACCCTGGGG TCATTTAAGG GCTGGTGTGG TCATTTACTT 22140
 AATCATCTGT GACCAGCAAG GGCCTTGTTT TCAGTAAAGC TCGGAAGCTT CCTTGGCTCT 22200
 TTATCAATCA TAACAAACAG CTAGAATTTA TTGAGAGCCT TCTCTTTGCC AAGTGCTTCT 22260
 ACTTGCTAAC TTAACTTCC TCCACCCTCA AGCCCTCTAC CCATTTTAC AGATGAGGAA 22320
 ACTGATGCTC AAGGTTGAGG AGTTGTCAAA GAGTACACAC TGGCCAGGAT TACGGAACCA 22380
 TCTTCTGCCA CTA CTGCTT TCTCTTGTTG GATATGGACG CTGTGGTTTT A TACTCTACA 22440
 CAGTTTAAAA ATGGTCGAAG TTCTCAATTT AGGGCAACTT TGAAAGGCTA AAGTGCTGTG 22500
 TGAGTATAGT TTTTATAATG ACAAATTCC AGAAGAGGAG ACTAAGTGAA TAGTTGCTGG 22560

ATGTCAGAGC TAATGTTGCT AGGAGGGAGG CCCATGTCCT GGGACCGTCT GGTCTGTC TC 22620
 AGGGGCAGTG GCAACTGTGA GGATCCAACC ATGTGTGCAG AGTGGCCCCA ATATGGACAC 22680
 ATTGTGACAA TTTCTGAGC TATAACCATG TAAGATGTAA CCTTTGGTGG TAATTGAGTG 22740
 ATAGGGACAT GAAAACTTTC TGGCTTATTA TTGTTGTTTG TTTGTTTCTA TTAATTCTCT 22800
 TAAGTACCTC AGAAAAAAG TGCTACTTAA TTCCATTGTG TCAAGATGAC CCAGTCTCAG 22860
 ATCAAGAGCC ACATTCTGCC CAAGCAGTTC ACACCATGCA ATTCAGGAC CTAGGAGGGA 22920
 ACAGTGTCTA GCAGAGAGAC CAGATTTTAA TGCCAGTCAG ATGTAAGCTG AGACTCTCTT 22980
 TCCCTTTTTA TGGAAGTGTT AACTAAGGG TTGGATGTTT ATACCCCAAT CTCAGGGCTG 23040
 TAGTTAGGGA CCCAGAGCAA GTTTCTCAA TTCTGTAACC TTTTCAGTTC CTAGCTGTCA 23100
 GGTAGCTATG TGAAGTGTAC CCATCTCTAG AAGCCAGTAA GAGAATCCAG TAGAACCTGA 23160
 TGGCCTAAAA TTGATGTCCA GGTCTTACAG AGTAAAGAGA GAGAGCTGAC TTCAGCAAAT 23220
 TGTCTCTGA TATCTACACA TGTGTGTACC TGAAAACACA CATCCCACTA ATAAAATATA 23280
 TTAATGTAAA CAAAAAATT AAAACTTTTT TAATAAAAGA AGAGGATCTA GCGAGAACAC 23340
 ATCCTGCCAA AAAACAAAA AAATTTTTTT TTAAGTTACA GGTAGTGGTG AACTGCCTAA 23400
 AATGAGTGCT GAGAACTAAA CTTGGGTCCT CTGGACAAAC AGCAAATTCT CTTAACCCTT 23460
 GAGCCATCTC TCCAGTCCTA GCCTTACCAC ACTCGTCACA GAAAGATATG TTGAGCTCAC 23520
 TCTAGACGAC TTATTGCTAG CATGAGTATC TGTCTAGTCC CATGTCTAAT CTTCATGATG 23580
 TAATCAGACC TACCCAGCAG ATAGCAAGGC AGCAGTAAAT GCTCTTTTTT ATTTTTTCTG 23640
 GACTTGGTCA TTTATTCTT CACTGTTATT ACTTTACTGA AGATTGGGC TGGCACTGGT 23700
 GATAAACTGA TAGGTATACC CAGGTGGTCT CTGCCTGTAT TTGTTTCTCC TCTATTGCTA 23760
 TGACAAAACG CCATGACCAA GACAACTTAA AAAAAAGAA AGCATTTAAT TGGGCTTATG 23820
 GTTTCAGGGG GCTCCAGTCC CTGACGATGG AGCAAAGGCA TAGCATCAGC AACAAGTAAG 23880
 AATTCACATC TTGATCCATA AGCACAAGGC AGAGAGCACA CTGGGAATAG CACCAGTCTT 23940
 TTGAACTTC AAAACCTGCC TCCAGTGACA TACCCCTTCC AACAGGCCAC ACCCCAATCC 24000
 TTCCCAAGCC ATTTACCAA CCATTCAAAA TATATTCACA ATATATGAGC CTCATGGTGT 24060
 TCTCATTACC TGAGACCACT AAAGGGCTTC GTATTTCTTA TCACATGGAA TCCTCCCATC 24120
 ATGTCTTTTA TAACTTAGAG TAGGCCTATT CCATGTAGAC TCCTCTACCA GATCCATCTC 24180
 GGAGCTCCAG CAATGCAGTC ATGTGACTGA GCGTCTCTGC CAGCCTTTGC TCTGAACTGC 24240
 ACATTCTGCC TCCACAGTGA CCAGAGCTGC AGACAATGTA TACTTAGGTC CATGCCCTAA 24300

ACAATAGATC CTAGACACAG AAGTCCTCAG CCCATTTCTT CAGAGAAGAG CAGTACCTCC 24360
 TATGTTAATC TTAGTAGCAG TGGTGGTTGT TGTTTTTTCT TGGTTCCTGT CAGTCAGTAT 24420
 TTTGACCAGC TGAATAACAT TTCTTATTTT AGCCTTTTGC ATCCTCTGAG AGTAAGATCC 24480
 TCTTGGCTTC AGTTCTGGTC TCTTTACTGA TTTTGAGTAC AACTGAGCCA TGTTAGCTGG 24540
 AAGGCAGACA TTGAATGGAA AAGTAGAGCT AGCATGCCTG TCTCTCTCAC TCATTGTACC 24600
 CACCTCTGAC AGGGTATGTA AGGGTACCCG TCCCTCAACC CAGCCTCAGT CAGCCCATGA 24660
 CTCTGGATGG GCCAGTGTGG TTAGCCATTC ATGGGGGTTG CATGTCTTAA ATAAAAGGGC 24720
 ATGGAAGGAA GCCTCTTTGC CTATGATCCT CAACAAGGTT CACATCTGAA TGCCATTTGC 24780
 TGTTCCTCTGT CTGCTTGAAC CTAGAGAAGG AGAGGTTGTA GCATGGGGCT CTTACATGGG 24840
 AGATAGCAAG TGGGAAATGC AGACTTTAGA GCCAGGCAGG TTTGCATCTA TATGCCAGTT 24900
 GACCAAGTGC TGATTTGCCT TATTTTAGCC AAATTACTAT ACCTACCCTA GCATCCATCC 24960
 TGAATCCTT TAAATAGTGG CAATGGTAAC TGGGCGTGTG ACCCTCTTGG CAACATTCCA 25020
 GCTGCACAAG GAGCCTGTGA CTCCTGCTTC TCCTTTTAGG GCTTTATCTG ATCTTGCTCT 25080
 TTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTAA TCTTTGTGGG 25140
 GCTTACCCAA AGTTGGGTAA GTCCAAAGTT GGGACTTCTG TATTAGAACT AGGATGGTTG 25200
 GGACAAGATA ATAGCTGAGC AGATACACAG TGGATATAGT GAACAGAACT GTATACTTGC 25260
 ATTTGGACTG CCTAAGCCAG TCTAGCAGGT TGTGTGGCT GCTTCCCTGC CCAATCACCA 25320
 ATAGACAAGT CTAAGGAGC CAAGGTCTGA CTGGGCTTCT ACCTGGCAAG ACACATCTGC 25380
 CAACCCAGCA TGGCCGTCTT AGGTTGTTTG TTTGGGGATT TGAGGAAGGG GTGAGAGTTT 25440
 ATTTGGCTAT TTGCTTATTT GGTAAATTTA TTAGTATTCT TGTTTGGTTG ATTGTTGTTG 25500
 TTGTTGTTTT TGAAACAAGG TTTTACTGTG TAGCCCAGGC TGGCCTCAAA CTCTCCTGCT 25560
 TCAGTCTCCA GAGTGCCAGA GTTAGATGCA TGTAATCCCA TCACTAGTGG AAGCCTTACT 25620
 TTTGAAGAGT GTAGCTCAGT TAGAGGTATG TAATGCCATA GGCTGAAGCA GCCCTAGAGA 25680
 CCAGTCACCA AGGGAGAAGG TTGGGGCTAC CATGTGACAG AGGAGCTGTG TCAGCCTGGC 25740
 CACCTGTGCA GTGGTGTAAG TACTACAAGA CTCCACTGAA ATCTGAGGCC CAGGTCTGCT 25800
 GTTATGTTTC CCAGGGAGGC ATGCAGAGAA AAAGTGTTT CCCTAATACT GCTCAAGTTT 25860
 AAAACAAACA AACAAACAAA CAAAAACAT GGTGGTACTT GCCTTTCATA CCAGTACTCA 25920
 GACAGCAGAG GTAGGTGAAT CTCTGTAAGT TCAAAGCTAG CACTATGTTC AAGGCCTGCC 25980
 AGGGCTGCAT AGTGAGACCC TGTCTAAAAA AGAAAAATGA AACTGAACCC TGAAGTTGTA 26040

GAAACTGCTC AGATTTTCAGT GAGTTCTTTT GGACTAACTG AATGAGCTTG TTTCAGCGCC 26100
 TTATTTTTTTC TCATGTGGAG CTGGCACATG AGCAAGACTA TCCCCAGGCT TTGCCACTAC 26160
 AGGATCACCA TTGTGGATAG GTCATACTGT TGGTCTGTGA TTTTCCTCAC TTAATTTTCA 26220
 CAACAATCTC AGAAGTGCTG TCATTATCTC CTATAATTCT TCAGAGTCAG AAAATGAGGT 26280
 ACAAAGAGGT AAAAGAAGGA AGATCACCTA ACTATTAGGA AGTAAAACTG GGATCCAAAG 26340
 ATGGGTGACC TTTTCTTCTA GTGTAATTTG CCTTCTGACG TTGTAAGGCC AGGGCACAGC 26400
 AAAGGAGACA GAAGCAGAAG TGTGAGCCCT TAGAATGCTA AAAAGAAAAA GAAAGTTAGA 26460
 GTGGGGAAAG ATCTAGACTA GAACAGTTAG ACTTGGTCTG TCTTCTGAAT TCTAGCTTTG 26520
 GAGCCCCCGC AAAGACTGCA TGTATATAC AGCATAGAGT TAAAAGGAGC ACAGGTTTCT 26580
 GCTTAAGAAA GAATGTGAGC TTAATTCAAT AACATTCAAT AGTATATATA GCTTCTTTTT 26640
 ATATTTTACA CTTATTTATC TTGTGTGCAT GTATATGTGA GTATATACAC ATGCCAACTG 26700
 CACACATGTG GAGATCAAAG AGCAGTTTAT GGAAATCAGT TCTCTCCTCC TACCATGTAA 26760
 GACCCTGGGA TCAAAGTCAG ATCATCAGGC ATCAGCAGGA GCCTTCTCGC TGGTCTCCAT 26820
 ATGCAGTTTC CTAAAGAACA AGGTTATCCA AGGGCTCTCT CACCACAGGT GATCACAGTT 26880
 ACATCACAGT TAGCAAGGCC AGAAGAATGC AAAGAATGTC TTTATTTCTT TCCTGGAGCC 26940
 TGGCTCCTGC CCTCCTAAAC TTCTTAAATT TTGTTTAATA TTTACATCTC TTCTAAGATG 27000
 TAAGTACTTG TGATGTCTTT AAATTTTACA ACACCCATGT GTTCCCTGGT TTACTACTACA 27060
 AGTAGGGCAG CATCTCTTAA ATAATGTTGT TCTAGAAGGA AGAGAGCTCA GATACAAGTA 27120
 GCAACCTGGA TAGGAATAGC AATTCCAGCT ATTGGATACT CACTGGATAT AGTTCTAAAC 27180
 AGTCTAATCA GCAGTTGTGT GATCAGTGGG CACTTAGGGC TGAATGGTAG AAGAGTAGCT 27240
 CTCATGCCAG GAAATGCACC AAATCAACCA GAGCAAGCAC AGACAATGGA GGAGAGACAG 27300
 GTGGCTTGCC CCAAGACCCC CCAGGAGCCT AAGATGGCAA TATTGTCGTT TTGAATACAT 27360
 TGTGCAGGCA CTTGGCCTCT GGGAGGGAGG AAAACAATTA GCTTAGCATC AAATCATGAA 27420
 CTCTGACAAC TGCTCTATCT TATATAAGAT CTCCTTACAT AAGGATGCAG AGAGAGCATC 27480
 CTCATTAAAA CACCTCAAGG GGTTTCACT GATTTTCTAG AAGCAGAGCT TCTCTCCCAA 27540
 CAAATACATC AGGACTGGCT ATAGACACTT TTTTCTTCAA TAGGCTAAAA AGATCCCACA 27600
 TTCCTCCAGG AGACAAACCT CAGAACAGCC ACAGAGGAAC TGGGCTCCAT GGTATAGGTG 27660
 GGGCATCTAA GGTCCCAGAG CCCACCTCCA TCCAGACTCA GGGAGAGAAC AGGCAAGCCA 27720
 AATCTGCTGG CTCTCAATTT GGTTTACATA ACTCCTGACT CCTCAAGTCC CTGGAACTG 27780

AGGCCAATTC CCTGGAAGAT CATCTGTTC TCTCCTGTTT TTTCAAGAA AGAGGCAGG 27840
 TGATCACTGG CTCCGAAGAC TGTGTGAGAG TGTCCCACTT CCTTCTTCCA CGAACTGAGT 27900
 GTCTGCCGTC ATGGCTGTTG TTTAGGAAGG TTCTGTTTGA ACTCTCATAA CTCCATATAT 27960
 GTTGACCTTG TATTATAAAG AACTTTACTT ATCTTATGTG TACCCTCCTC TGTTTCCAAA 28020
 AGAAAATGGA GGA CTGTGCAG CAAAGGAAAT AAGTAAGGTG AATACATTAG GAGAAGTGAG 28080
 AGACTGGGAA GGGAGGCAGA CAGAAGGTGA GCTCCAGTA TCTGTGTGCA GAGTAGGCAC 28140
 CAGACTCTCT ACTGCAGTAT CGCAACAGCA GAAGCAATCC TACCTCAGAG AGTTGAGGGG 28200
 GAAGGTAAGA AGGCACATTT TTTTAAATAA TAACAACTT GACTGAAAGT TGAAAGATGT 28260
 GTTCCTAGTA CTAAGAACAG TTTCTCATGT GAGGTTGCCT TTAGGGGCAC TGCATACACT 28320
 TGTAGCAATG AAAAAAGATG TTTATAGGCT CTGTCTTAAG GTAAACTTGG TGAGAATGGA 28380
 GGGTAACTAA AACAACTTAA GGAAGGCCAT GAGTCTGGGG AGCACTAGCT CTTTTGGAGC 28440
 CTCAGTGTGT CCTGGGTAAA GTTGAGCAT CTTGTGGTG GCAGCTCGAT TGGTGCACTA 28500
 AGTGCAAATG TGCACCAAGT TCTGGACTCA CTCTTCTCGG ACACATAGAC TGAGTGTTGGC 28560
 TCATATCTGT AATCACAGCA CTCAAGAGGT GGAAGTAGAA GGATCAGAAG CACAAGATGG 28620
 AACAACTCA GCTCTATAGA CAGTTTAAGG CTATCCTGGG CTACCTGAGA CCCTGTCTAT 28680
 AAGCAAATGA CTAAACAAAC AGACAACACA CTTAATTTTT TTATAGCAAC CACTTTGAAG 28740
 TGGGAGGGGT CTGATAGGGT CTCTATTGTT CACAGCAAGT GCACAAGGTC AAGAGTAGCT 28800
 AGGCAGATGA AGAAGAGGCC AAGACACCTG AACAGTATCT TTCCCATGGG TTCGGAGGAG 28860
 CCACGTGCCA CCTTCACAGT CAGCATTGTC TGTGCGAGTA GCTCTGGCAG CATCAGTGCC 28920
 CAAACAACGG CTGATACGAG TCCCAGATG CAAGAGGAAA TAGTTGTCTG TAATTGCCTG 28980
 TTTTAAGTAG AGTGGTCAGG AGGCTACAGC CTCCTCATCG GGCTACATGT GGCATATGCA 29040
 GGCTTGCTCA TCAGACCTTG TATTTACTGT TTTCACCTTA ATGGAGAATG GGAGAGGCAA 29100
 ACAAAGCCCA GGGACTTTGT GGAAGCTGAC TAGAAGCCTC TGGGACTCCA GGGACTGCCA 29160
 ATCTGCTAAA GAAGAAGCTA AGAAAGAAAA TGAGCTCCTC TGCATGGGTC TCCCATGAT 29220
 GGAAACAGAA GGCCACATGG CACAGTGTA AATAGAGCCCT GCTGCACTGC TCTTACTGTG 29280
 GTGAATGAAG AAGAGGCAAC TAGCCAGGAG GGCAGGACCA CTACTACTGT TTTGCTGGCT 29340
 GGTTCCTCCC AAGTGAGCAG CCTTCCCTGG GGACAGACCT TAGCTCTAAG ACAGACGTGG 29400
 CTTCTTCGGA GCAAGTCAAA CCTCAACATC GAAGAATCCT TGTCTTGTC GTTTTAGCTT 29460
 TAACAAGAAT AGAACAAGCT TCTGGAACAG GACACAGTGG AGTCAGGAGA AGCGGCCTTA 29520

AGTGAAGACA CAGCTGTGGG GTTTCAGAC TCGCACTGCA GGGAGGCGTC ATCCAGTGGG 29580
 AGCGGCCAGC CTCGCTGTAG ACTTCCAACA CTAACGAATC GGGAACTCCA TGCTGAACAG 29640
 GATTTAGTTA GAGGGTCCCT GTGCCAGCAG ATGGATGTAT TTTTCTTGAA AGACCAAGGT 29700
 GCCAGAACTC TTCATGATTA CGTACTGGA GCAAGGTCCT TTTTGTGGT TTGTGAAGTT 29760
 GAGCGTCAGG ACTGCAGGAT TCTCTTGCTC TTTCTTACTC TTATTTTTC CAGGTCAGAA 29820
 CCAGAGCTTG GAGCAGGGAG GAAAATCCTG CTGAATGAGC AAGTCTTTC TAAAAAGCT 29880
 CTTCAAGTCC AAAAAGACTT CAGTGGACTT AGGAGAAAGA AATTTAATAC ATTGCCATAG 29940
 AATCGTTGTT AACCAAGTTA AAGCAAAGCC CACAGCATCT TTGTCTTATA AAAGAAAGCA 30000
 AAGAGGAGAT GGAAAAAAG AAATAATGCT TAGGAAATCC AAACCAAACA ATGAAGACTA 30060
 ACGAAGGAAA ACTAAAGATC ACTTCAAAGA ATGTGAAGAT TCCCTCCTAA TAAGATTTTT 30120
 CAATTTTCAA ACCTAAGCTT CAGGTGGGAG GACCTTTTCA GTTTTTTTTT TTTTCAAGTA 30180
 TGCTGTTAAG TGGCATTCCC CAAAATGTTG GCCCTGTGTA GGATTGGCTG CTTCCACAT 30240
 AAGGAGCAGT CAGATACCCT GCAAGACCCA GGAAGTGGG GAGCTTTAAC CATGGGAAGC 30300
 TGAGAGGCTT GCCAGACTGC TCCTTGACCT GAGCTTGAAC CTGAGTCCTA ACTGCTAGCA 30360
 AACTGAAACA AGCCCAGCCT CCAGGAGAAG AAAGTGGGCG GAACTAGAGC AGTCCTAGCC 30420
 AGAAACTAT GCTCCTTTCA CCACTGGCTC TGTCTTTACA TCCCTGGGAG GGAAGCCTGG 30480
 GTTGGGCTTC AAGATCGCCT GCTCAGACCA TCCCTCTCAC TTGCTAGCCC CTTCCAGGCC 30540
 CACGCAGAGG CACTAGTGCC TATGAGAGGT CAGTTTGCAT CTGTTGTGGA CAAGACAGGG 30600
 AATTCCTTGA CATTTTAAAT ATTTATTTAT CTTTGTTAGT GTGTATGTAT ACACACACAC 30660
 ACACACACAC ACACACATAT ATGCACAAAT GTACCAACAA AAAGTTATGG AGCTTGTGGG 30720
 GGGAGTCAGT TTTTTTCCTT TCACCATGAG GATTCCCAGA ATTGAACTCA GGTCAACAGA 30780
 CTAGAAGCAA GCATCCTCAC CAACTCAGCC TTCTCACTAT ACCTTGCATA GAGTTTCTCA 30840
 ACTTTTGCCT AAGCTCAGAC TGGTAGTTTT TTGTTTTTGT TTTTAAAGAT TTATTTATTT 30900
 ATTATATGTA AGCACACCAG ACACACCAGA AGAGGGCATC TTATGTCATT ACAGATGGTT 30960
 GTGAGCCACC ATGTGGTTGC TGGGATTTGA ACTCAGGTCC TTTGGAAGTG CTCTTAACCA 31020
 CTGAGCCATC TCTCCAGCCC CAGACTGGTA GTTTTTAAAA GCACCAGAAG TTCTGAGCTT 31080
 CCATCTTCCT TACTCAGTGA GTTTAAGAAG CACCTGCCTA GGCATGATAT TCTCCAGGGC 31140
 AGGCCATTTG GGCAGGCCAT TCTGTACATC TGAGCCTGTG AAAGACTGGC TTGTTCATTG 31200
 ACCCAAGAG ACACCTGGCT GCACACTGAC CACCCTTTC TGTTCATTG TGTCACCTTC 31260

TGTGCTTAT TCTTATGAAC GCATTTGAAT CCACTGACTT CACTGGGCTG GGATCGAAG 31320
 TAAGGCCACG TGCCTTTTAC TCATCATAGA AAACAACATAT AGGCCTCCTA GCCTCCTGCT 31380
 TAGCCTTGGA CATTCAATTCT CTCCCTAGTT TTGCTCACAA CATGGTAGAA TCTGAGACCC 31440
 AAAAGGACGC CCTTTATTTTCTCAGCCAAC TAGTAGTGTG GTTCCTGGGA GGAGACACTG 31500
 CTGGTCTCCC TTGCCACTAT AGTAAAACCC AAGAGGTGCA ACAACCCCGG AAGAGCTTGC 31560
 TTCCTACCTT CCCCAAATCC GTGGGAAAGT TTGCCATCCT GTCCCAAGGG TTTCAGCCTT 31620
 TATTTAACTC AGCCTTAGTC CTATGGCCAG ATGCCTTGTT CACCCCTATC ATGGAGCCTG 31680
 GACAGTGAAG GGCCCCATCA GAAGTTTTAT GTTCTGCTGC CCACAGCTGC TCTCCTGTGT 31740
 GGTCCTCAGCC TAAGTTTCTA GAAATAAAAA GCTCTCTCAC TCTCACACAT GTTCATTCTC 31800
 TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC 31860
 TCCTTCCTTC CTTCTTCCTT TCTCTCTCTC CTCTCCTCCC TTCCAACCTC TTCCCTCCTT 31920
 CCTTCTTTTT ATTTCTTTTT GTGAAGCAGA GTCTCTTTAT GTAGACCAGG CTGGCCTCGG 31980
 ATTCATAAGA GATCTGCCTG TCTTTGCTTC CCGAGTGCTG GAATTAAAGG TGTGTACAAC 32040
 CACACTCAGA ACTCTTCCAT TTCTACCTAA AGAAGACCTG TTTGTCCTTT GTCAAGCTGA 32100
 GAGCCTTTTCG TCTCCCTAGG TCCCTTTCAA AACTTTATTC CTGTGGCAAT GGCCTAGAAG 32160
 CCAATCCCTT TGAGAGGACC CACTAGCAGT CAGTGCTTCT GTTCCATGTA GCAGCTGCCA 32220
 CCAGAGTGGC TTCCATTCTT GCTGGCTGAC TTCCCACTGA GGGGGGCCTA CAGAGCTTCG 32280
 TATGTGCCCC AGGCTGGCAG AGAGGGCAGC AAGGAAGGCT CTGTTCTGGC AAGGCTTATG 32340
 GTATAGGAAG TATCTAGGAA ATACTGTTGC TCTTCAGGGT GCTGACAAGA TAGGAGCTCT 32400
 TTCTTGCTTC CCGGGGATTT GGACCCCTAG TTTCAGTAGA GCTGGTCTTT GTTGAAGTGC 32460
 TCTGCCTGGA TGTCTCTGCT TGTAGGTCTT TTGTTCTGCT TCTCTTGGGA ATTCTTCTGC 32520
 TTGCTTTCTG GCTGGAGGTA CTGGTACAGC TGCACTAGCC TCTATACTCA TTGTACACAC 32580
 TCCCCTAGCT TGTGGGCCTC AGTTGAGTCA CACATCCCTT CATGAGCTGG AACTGCCAG 32640
 CATGGATATC TGTTCAAGCA CTAAGAGGAT AGGCCTCCCT TAGCACTGTC AGGTCCAATC 32700
 TTTCTCTAGA GATTGGGTCT GCTTTTCCCT GCAGCCCTG GATGGCACAT CATTAGAAAG 32760
 AAGGACATGC CTTCCAGTGC TGCCTCTGTT TCTGCTTACA GGGATAAGTA TGTTTATTCA 32820
 TTCATACTGA ACTTTGTACT TGTAGGCACC TCCATGCCTG TAGACATGCC TGATGGCTTG 32880
 ACTTCTCTGA GAAACACATC ACTGTCCTAG GTAGATTTTA GAACTTAAGA GAATGGTACC 32940
 CACCTTGCTC CATCCCTACC TCTCCACTCC TTGGCTTTTC TTTGAATATT TTAATTACCT 33000

GTCCATCCTA AGGTCACACA CAGTCTAATG TCTGGACACA GTTCCTCCCA CCTCTCTAGA 33080
 GTCCATAAAT ACCTAGGAAG CCAGTACAGC TTTACAAAGA AGACTGCTTC TTCTGACTGG 33120
 CCCTTATGGG CCTAATACAT ACCAAATCTC TCAAACACAG TGTAGTGTGA GAATCTAATA 33180
 AGATCATATG AAGAATGTTT AGAGCAGATG TACTTCATAA ATATTAGTTT CCTACAGAAC 33240
 GTCTGTCACT CAGACCCTCT GCTTTCTCTC AGTTGGGCTG CATTTCTCTC TCATGTCTGT 33300
 CAGTACTTAG TTCCCTGGCC CGTCTGTATC CATCTGTTGT CATATCGTAT TGCCCTCCCT 33360
 TGCCCATAT TCATCCCTCA AACCTTTCTG GAAAGATCCA GCTTTGGACC AGCTTGCTT 33420
 TCTTCTTCAT ACTACTGTCA AGGCTGCAGA GGGTTGTTCA CTAATCCTAG CTA CTACTGAGTG 33480
 CTCTTTGGTG GTCCTCCTGC GTGGCCCCAT CTAGGTCTTC GTCTTGTCCT CCAAAGATT 33540
 GACTGCAACC TTCATCCTTT CCTCAAATTT CTAATCTCTC AAATCTCCA TTCTTTGCAG 33600
 ATAATTTGAC TTTCTAGTTC TCAGGAGGAC AGAAGCCATG CTAGAAAGTT CTAAACCCTC 33660
 CTTACCTGGC CTACAGACCT GGCTCTGTCC CTGCTCACC CTCCCATCTC TAGAGAAGGT 33720
 CTTCCATTGT GTGTTGGATT CCAGTCCCTG GCCATCTCAG AAGCAACACT GTTACTCCA 33780
 TCTCTTGGGC TCCCTCTCCT ACTATAGTCA ATTCTGCTTT AAAATGTCAC TACTTATATG 33840
 TACACCTTTC ACTCCCTTAC TCACTGTGCT GTCCCACTGT AGTCTCTGCT GTCTCCTCCC 33900
 TACAGCCATC TAAAGCCACT TTGACCTCTG TTTCTTGCT TCTCACTTTC CAATCTGTCT 33960
 CCTACCCACC TCAGCTCCCA CTACTACTTC CCTCCAGCCC TTTCTGCCAG ATCCAGTGGG 34020
 GTCCCTGTTT GGGACACACA CTCCTCTCCT ATGTGGCATT TTAGGAGGGT ATAACAACT 34080
 GACTTGCTC TTCCTTCCTT AAGAATTCCC CCTTAGCTTC TTCAAGACAT AAATCAAGAC 34140
 CCACAGCCAC CCTTCTTGGT CTCTGCTCCC AGATCTCTCA TGGAGGTGTT CTTTGGACTC 34200
 CACTAGGATC TTCTTCCTCC CATGCACTCT CTCAAGACGA TCTCACCAC TGCAGCTAGC 34260
 TCTCATCTTG CCAGTTGAAG CCTGCACATT CACTTGGAAC ACACATACAG CAGCCTTCTG 34320
 GCCATCCCCA CAAAAACAA AGAAACCAAC AGCTCCAAAT AGGACCCAAA CTCACGCCC 34380
 AAGCTTACCA TCCCGCATCA CCTGCAGGAG TGGCCTCACC ATCTGTCCCA CCATCTGAAG 34440
 CAGAGAACT GTGACACCTC CATTCCCCTG CATATCCAGA CCAGCAAAGT TCCATAATGT 34500
 TCTTAGCAAT GGACAAAGAG AGTGAGTTTG AGTAAAACT CTAGTTCTAT TGTGCTGTGG 34560
 ACAAATTCCT TAAGGATTTG TTTGTATGAG TTTGTTTGTG TGTGTGTGTG TGTGTGTACA 34620
 CATGTGTGTT TGTGTATATG TGGGTATATG TGTACGTATA GAGATGTTCT TGTATGTGGA 34680
 AGCCAAACAA CCTCAGGGGT AGTTCCTCAG GTGTTGTCCA CTGCTTCTCG TTGTTATTGT 34740

CTCTCACTGT TCTGGGTTTA AGAAAGCTAG ACTGGCTGGC TACTGAGTCC CAGGATCTGC 34800
 TTATCTCTGC CTCCCCAACA CTATTACAGG CATGCTCACA GATGCACATC ATACCTAGCT 34860
 TTTAAAAACA TGAATTTGGG GAATCAAATT CAGGTCTTTT TGCTTGAATG GCAAGTACTT 34920
 TACCGACTAA GCTATCTCCT TAACCTCTCT CAACTGAGCT ATCTCCAAAG GCATACAGAC 34980
 ACACACACAC CTCTCAACAG GATCTCAATA TGTAGCCTAG GTTGTCTTAA AACTCTAAC 35040
 CCTTCTGTCT CAGAATCTTG AGTACAAAA CTGTGGGTGT TCATTACTGA ACTCAGTTAA 35100
 ATTCTTAATC TTTATCAGCC CCAAGCTCTG CATCCATTAA ATGGAAATTA TAACACCTAA 35160
 TTCAAGTGGT CATCAGGATA AAGGAAAGCC TTCTTCACTT GGTGTGTGTT TGATAATAAA 35220
 AGTATTTAAA TAAATAAATA TTCAATAACT GAGTGCCCTT CTGTCCCTCT CTCCACCAAT 35280
 CGGACTTGTC TTGTTGTAA ATTGCTGTTT CTATAGTTTT CTGACCTTGA AGCCCTCCCC 35340
 CTCAAGATCA CACTTACCAG TGTTCCTTG ACTGAGGACC ACAGTGCCTG TTTCATCCCT 35400
 CCTTTTTTTA CTTTTGGGGC TAGGAGGCAG ATTCTAGAGT CCCATTACA GGGTTTGATG 35460
 TGTCTTCTCT CTAAGCTGTC TCTAGATGCC CCCATCTCCA CAACCCTGCC TGAGACCCAG 35520
 GCCTAATCTT TTTAGTCTGC CATGATGGCC TTGCCCAAAG CCCTTCTCCT GCAGGCTTGC 35580
 CTTCAGTTTA GCCCATCTCT CGCTGACCAC CAGGTGTCTT GTCTCTGAC ACCTGCTGTG 35640
 CTTTCCTTTT CTTTTCTTTT CTTTTCTTTT CTTTTTTTCT CTTCTCTTCT CTTCTCTTTT 35700
 CTTCTTTCTT TCTTTCTTTC TTTTTCTTT CTTTCTTTCT TTCTTTCTTT CTTTTCTTT 35760
 CTTTTTACCA TGTGAATTCC TCATACTCTT ACATGCAGCT TGCTGCATTA GCTCAGCCCT 35820
 CCTGCCTCCC TGAAGCAGCC TGATATCGTC CTCCTTGATC TCATTCCTCC CCCCCCAT 35880
 GTTCCTCTCC CCCCCCCCCC TCCACGATAC AGAGGAGGAA AGCATTGGG AGTGGTTGAG 35940
 AAACTGAATC TCGGTACAGC GACCAGTAGG ATAGACTGAG ACATTCAGCA AAGACCAACT 36000
 CTAAGTGAACC CAGGAGCCAA AACTCTGCA AAACAAGAAA AATGTAACAC AAGAGTGGGG 36060
 GCATGCTAGT CTTTACTCAA AATCAAAGTA GAGCTACCTT GTCTCGAAGA ATCTAGAAAA 36120
 TGCCAATAAA GTGGAGAATC CTCCCACTGG GCTGTTTCTC TCTCTCTCTC TCTCTCTCTC 36180
 TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCACACAC ACACACACAC ACACACACAC 36240
 ACACACACAC ACACACACAC ACACACACAC ACACACACAC ACGTCTCTCC CAACCTTTTT 36300
 GTTTTGGTTT GGTTCGGTTT GGTTCCTGTT TTTCGAGACA GGTTCCTAT GTATAGCCCT 36360
 GACTGTCCCTG GAACTCACTT TGTAAGCCAG GCTGGCCTTG AACTCAGAAA TCCGCCTGCC 36420
 TCTGCCTCCT GAGTGCTGGG ATTAAAGGCA TGCGCCACCA CCACCCGCTC TCTCCCAACC 36480

TTTTGTGAT CTATTTTTTT GTGGTTTCCT TAGCATGCGA TCAAATGTAT GAGCTGCTTT 36510
 ATCTGCCCAC CCCACCATGG CTACCTGCTC TCCCACATGG ACTGCAGTGG GACCTGTCAT 36600
 GCTTCCTGAC TTTTGCTACC AATGCTGGTC TTATTACCAA TGCAGTAGTG AACTGAGGC 36660
 AAACTGTTTG GCAGTGAAAC CTTTCTCTAA GCCACAAATC CATAGCTTAA AATATTGAGG 36720
 CAGAAGATGC AAAATTTTCT AAGAGTGTAG GTTTTTCTGT TTGTTCAATTT GTTTTTAGTG 36780
 GACAAAATCA ATACACTGCC TCAGCTAGAA AGAAAGAAGT GAGGCAAAAG GTCATAGTTG 36840
 TGATTAAATG TTGTTGTAAT TGATCTGCTA TACAGTGGGT TTTTTTTTTT GTTTTGTTTT 36900
 GTTTTTTTTT TTTTTTTTTT TGCTTTGTTT GGGGATGTTT CCTTTTGACA CAGAATCTCA 36960
 GGAGGTAGCC CAGGATGGCC CTGAACTTTA AACCTTCTGC CTCAGCATCC TAAATGCTAG 37020
 AACCAAACAC ATGTACAACC ACACCTATCT ACTTATGTAC TAATTATACC AAATAATGGA 37080
 TTTGCGTTGC CCTTTCTATA CACGTGTACT TATACTTCGA TGGTCATGCC CATCACTGTG 37140
 TCTTGTTCCC ACTCCCCTGG CCCTTCCAAA ATAGTTCCTC TCCTCTCCTC TCTCTTTTTC 37200
 ATCTAGATTC CATGCATGAG ACAGAATATA TTTGTCAGTC TAGGTCCAAC TTATTTTACA 37260
 TAACAAATGT CAAATTTTCA AATGACAATT TTTAATTCT TGTTCCTTAT TTCATTTTCC 37320
 TGTGCTTATA CATGTGTGGT GCATGTTTGG TGGGTGTGTG CATGCAGAGG CTTGGCAGTC 37380
 ACCCTCAGCT GCTTTTCCAC GTTTCTCTCT GAGGCACAGT CTCCCATCAC GTCCAGGGCT 37440
 CACTAGTATG GCGAGTCTTT CAAGCCGGCT TGCACTAGAG ATCCCTCTT TCCTCTCTGG 37500
 GATAGGAATT CTCGGCATGT GTGTGAGTTC TGGGGAGCCA CCTCTGGTCC TCATACTTAT 37560
 GCAAAAAGTG TTTTAACCAC TTGGCATTCT CCCAGCTCT CATTCCTTTT TATTGCTGAA 37620
 TAAAACTCCA CTGTGCGTAT GTACCACATT TTCTGTATCC CTTCTTCCCT TGATGGGATC 37680
 TAGACTGGTT CTGTAGAAGT GCCATGAAAA CTGCTTTGGT ACAGATCGAT GTCTGTGTTG 37740
 TGCTGACTTT GTACTCCCTT CAGACAGATG TCCAGAGGTG GTAGAACTGG ATCATAGGAT 37800
 AGTGCTATTT TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCTCTC 37860
 TCTCTCTCCC TCCCTCCCTC CCTCCCTCCC TCCCTCCCTC CCTCTCTTTC TTTCTTCTTT 37920
 TTGGAGAAGC CTCCACACTG ATTTCCATAG TAGCTGAACT AAATTCTTTT TAATTTAACT 37980
 GAAATAGAGG CCTGCTTAGA GCCAAGGTAT AATCTGTAAG AAAAGCCTTT GACTCCAGTG 38040
 AAGTTCCTGG CTTTGTTGTG GTAAAGAAGC ATTTGTTTCT AGTTTGAGTG TTCATCTGGG 38100
 TCAGTAAGAG GACAGACCAT TCCCAAGAGT GTGCTTTGCT CTGAGGGAGA GAAAAATTGT 38160
 CCAGTATCTA ATGGTGCAAA TCATTAGTTG TGTTAATAAC CCTACAGGGG AAAAAATCAT 38220

AATAACGTAT CCCCCTTTCA TGTACTTAAT GTAGCTAAAT TTTCCCTAAT GAGTTAAAG 38280
TCCATGGAAT TTTTGGAGAT AGTAATTGGC TCCACATTGG AAATGCTCAA GCTCCCTGAG 38340
CCCTGGGCTC CAGTAAGACA GGTAGTAAAC CTGCCTGAGC CCTATGAAGC CCTGTGTTC A 38400
CCTGAGGTCT CCTTGCCAGA GTCCCAAAAG AAACCAGGAC TCAGCAGGTT GTCTTTTCAT 38460
CTTCACTTAC AGGGTCGCCA AGTCAGTGTG CTCCTAACCT AATTCTGAAC TCCCTTCTTT 38520
CCCCACAGT GTAATATTTA CCTAACTGGT TGTTCAAATC AAACCTTGAA CCTTGGTTCT 38580
CCTTTGTCAC CATATCCATC AGCTGTTGAT TCTACTTCTA AAACCTCACTG CCACTTAGAT 38640
CTCCCGAGTG TCCATTTCTC TTCATCCTGC CTACCTGCCC TCTGGTCTCC ACTCCCATT C 38700
CCTTAGGAAC AGCCCTGTGT AGGCTCCCAT TTCTCTCCTC TTCCTTCAGA CAGCACAGTA 38760
GCCAGCTAAA AGGGTCTTCC CAAACTGAAC TGTGGTGACA TCATCCACCC CTTTCTTAGA 38820
GCAGATCACG ATCCTTCCTT TGTAGTTAAA GCAAAGGTCA GAGTCCAATA TAGCCAAGAG 38880
TGCCATGCAT ATTTTGCCTG GCCTTCCTTC CACAGTGAAC CACCCACTGT GATCCTTGTA 38940
GCTATTGGAC CCAGCTTCTG GCCATCCTTC AGTTCCTGCA GCTCCTGCTC AGGACCTGTG 39000
TATATGCTGT GACTTCTGCC CAAGACAAC GTCTGTTCCTC TCACCTGTGG GCTCATCACC 39060
CCTCCTTCCC ACAGCCCTGT TCTGCCTGGG TCAGCTCATT TGCTTCTCTA GAAACATGGT 39120
AAGTTACATT TTGCCTCCCA TAAAGATTCC AAGAACCCTC TTCAGTTCAA ACTTCTACAT 39180
ATACCTGGCT GTTCCCTACA ATAGAGGTTT GCTCTTGTGA CAGACAGCAG GCTCCACAAA 39240
GGTGACCCCA CTGCTGTAGT GCTCCTGTGT CCTCTGCTAG CTCAGAGGCA GGTATAAATG 39300
TGCTTTCCGA GTACGAATTA CATGGCCAGT GCATAAGATG CCCCTGGTGG CAGTACCTGT 39360
GGATGGCAGA GCTCCAGTGA CAGGACTCTG CCACAGTGAG AGGTTGGCGT TTCCACAGCA 39420
CATCCACTGC AGGGAGCTGA CTTGTTTGAT TCTCCCTCCG TTGAGCCCAA ACTCACTGGT 39480
TTTTCTTTTC TCTTTGTTTA GCTGAAGCTT GCTTTTACTC TGGACCATGA GACTGGATTG 39540
CCTCAAGGAT GTCACATCTA TGAGTACCGA GACAGCAACA AGTAAGCCAC TCACTCAGGG 39600
GAAAGCATCG CCTACTTGCT GAGCAGCTGG GGCAGGCTTT CTGACTCGGG CTTCCCTTAC 39660
AGGGGTGCTC AACATTTTGG CTACATGTGA GAAAATGTCT GGCACACACA TACAAAATAT 39720
CCACCCCAA AGTCTCTTTT GACCTTAAAT ATAATAGAAA GGAACCTGTA TAGAGGGCTA 39780
GAGCAATGGC TCAAAGCACG TACTGTGAAA GTGTAAGGAC CTGAGTTTTA ACCCCAGAA 39840
CCCACATAAA GCCAGGTTCA ATAGCACAAG TCTGTAACCC CAGTATTCCT ACGGTGAAAT 39900
GTGAGAAAGA GAGAAGAGAG TCCTTGAAGG TCAGATAGCC TGGTATACAG AAAGCCCCTG 39960

TGCCAAACAC TGTGGAAGGT GAGAACCACA TTGAAGTTAT CCTCTGATTC CATATTCTT 40020
 CATGGCACAC ACTCATGAAC ATTTGCACAT AAATGTGTGT GTGTCACACC ATACATATAC 40080
 AATCATACAC GCATGCATGG AAATAAAGCA GGTGTAAAAG AGTTGGTTAA GAGTAAGAAG 40140
 TGTGTTAGGA AACCAGCCCT CCTTTTCAGG CCCCACCCTC CTCCCCACCC CTCCCCAGTA 40200
 CTCGCCCTT GCCTGCTTAT CTGAGTCAGC TGTGACTTTG GCCTTGTTG TGGTTCTGTA 40260
 GCCACCGACT CCCCACTTAC TACTTCTGTA GTGATCCTGT GGCTGTGTAG TTGGGAGGTG 40320
 GACACAGATG CAAAGTAGTG TAGCCCTGTT AGAAAATGGC CCCAGCATAA TTTTAAAGTA 40380
 CCTTTTCTCT CCTTCAAAGA CTGATTCTCT GAGTGTGTGT TGTGGTGTGG TGTGAGACAG 40440
 GGACAGTGGT GGTGAGTGAG GCAGAAGAAA TATGCCTTGA TAATGCTGGT GCTGGTGGTG 40500
 GTAGTAGTGG TGATGGTGGT GGTGGGGTGA TGGTGGTGGT GATGATGGTG ATGGTGGCAG 40560
 CAGCTCACAT TTGGGCACCT GCTCTGCATT AGACTCATGG GAACCAGTGT GTGCCATTCC 40620
 TACTTAACCC TCATCACAGC CTGAAGAGTG CTTTCATTAC TATGCACTGC AGAAGCTAAG 40680
 GCCTAGGGAA CTCTGCCAGC TCACTCTAAG TAATTTACAT ACACAGTCAA CTTTAACATG 40740
 TCTACAGTGG AGGAAGACTA GGTGGAAGAC AGTTGTTACC ACTCTGGGAA ACCATCCTCA 40800
 ATAACCAGTA GACCCAGCCT AGACTTGAGA ACAGTGTGTT CTGGTATCAT CATATAACTA 40860
 TCTAAACTAT GTAATCTCAC CCAGCTGAAG GAATAGGCAC CTGCCAGCAT AGCCAGCCAT 40920
 GACCTCCCAG AAGAACTCAC TGCTCAGATG TGAGTAGAAG ATAGGTCAGT GTTACCCTTG 40980
 TGACCACATC CACATGCAGG TTGCCTTCTG GGTATCATTG CAATGTCTGT ATCTTTAGGC 41040
 AGATGATGTA CTTATTATTG GACAACACTA ATTCCCCTT CATGAACCAT GGAGAAGGCC 41100
 ATCCAGTCAT CCTCAATGCC TCTATCTCCC ATCTTGTATG GGGCTACATC TAGAAGGCAT 41160
 CCCAGTGCTT CTAAAGCCAT TGTGTCAAAA ATACCATCTT GGTTCCTTAT TAAGCTCAGG 41220
 TCATCAGCAA AGTCAATCTG GGATTCTCTG AAGCCAAAGG CAAGATGGGA GAACTGAACA 41280
 GATTCCTGAG TGGCTGGGCA AGCTTTCTTA GAGACTAAGC ACATAACCCA TAAACAGTAC 41340
 AGCATACCGT TCTGCTTCCC TTCCCTTGCT CCACACTGTT CTCCATGCCT CGGCCCTCAC 41400
 TCTTAGCCTC CAACTGCTTG TCAGGATGCT CTGTTTACTG TTAGTCCTCT GCAGAACCCT 41460
 CTTGCCTTTC AGCCACCAGC CAGCCTCACA GGTCTGCACA CGGTACCTTC CAGAGCTTCC 41520
 CAGTGATACA AAGCCATCTT CCCAGGTCAT CCTGGTATAT TTGAGTTATT GGAACAACTG 41580
 TTTGTCCACA GACCCTATCC ATGCCACAT ACCACTTAGC GGCCTCTCTG TCCAGTACTT 41640
 ATCAGGAGAC TGGCAGGGCA GCCATAGGCC TCTCTCTGTA CAAGCCTGAC CACTGGGAAG 41700

GAATGGAGCA TCTGGGTAGG GACTCCCAGG CTGCACTTAC TTTTAAGTCA TTTTCAAGCAG 41760
TCTATGGGAA GCCTCAGTGC CAATGCCCTT TGGAGCCAAC TCCCCTTCTT TAGGGCCTGG 41820
CCTGTGTCTG GGCTCTACAC ACATGGGGTA ATGCTAGATG ACTCAAGACA TTCAATAGGA 41880
AGAGGGCTCC AAGACAGCTG CAGCATCAGA ACTGAGCAGC CACGTCTGGG ACTATGGCAG 41940
GGGATCGAAG TGTACCTTTC CCTGTGTACC AGCCTAGGCG GGGGGAGCAA GGGATTCTGG 42000
ACCAAGTCCC ATGTTTAATT AATTCATCCC TCTGTCTACT TGATTCTTCT CCTCTTCCTT 42060
CCCCCTGAGC AAGCTGATGA AATATTTCCC AGCAGCCCCT GACAACTTCA AACCAACATC 42120
AGCACTTGCC AGCACTTTTG AAATGGCACT TTCTGTCTGT GCTTAGAGCT ATTGCCAGTT 42180
CTGCAGACTA ACTGCAGTGT TACCTAAGAG CCACTCCTGA CAGAGGGTGA GCACCTCTAG 42240
GCCTCCCGCA AATACAGACG CTACCAGGTC AAAACAAAGA ATGATTTTCT TGTTCCTTGT 42300
AAAGCCCCAG GTTTGGAGAA AGAGAAGCTG AATCAACTCA GAGATAGGAA GGGCTTGCAG 42360
AGCTGGAGGC AGCAGAGCCA TAGAAGTGCC AAAAGTGACC TCATGGGAAC AGTTGGAGCT 42420
GGAGCATACA TGTGGAGTCA GCCACTCACA GTGCAGGGGT GGGCTTCTGT GACCCTCACA 42480
GCAGGTGGGG TTTGGTATCT CCATGACACC AACACTCCTG CTTCCAAGAC TGAGCTCTGA 42540
GATGATGTCT CCCCCTGTC TACCACACAG AGGGGTAGCC TTGGCTCGTC CTGTTCTGT 42600
TACCTAGCAT GAGACACCAA CAGCAGCAAC CAGAGTATGC TGGGTGCTAA AATACAGTGT 42660
TTGATTCCAC TTGGTTCCCC TAACAGAAGG TAAGAAACCA TACATGTTCT TACTTCACAG 42720
AAAGAAGAAC CTGTGATCTG AGAGATGCCC TTCCCAAGGT TGTATTTAAG AAGCAGACAA 42780
GCTTCTTCCA GGGTGCTGCT TCCTCTATGA GGTGCATAGC AGACTTGGGC CCAGCCTGTG 42840
GGTCTACAGA GATCTGATGC CAAGTTGCCT AGGAATCTGG GACAGGGAAG TCAGCAGGAC 42900
TAGGGTTGCT GCTGCCCCAT CAGGGTTTAT AGTACCTTTA TGTATTGTGT GCGGCACCTT 42960
CATAGTCGCT GTCTATATAC ATGTAATCTG TATGTCCAAG ATATTTATTA GGGGGGCTAA 43020
CTCAGCATCA TTTCTCAATG AAGTTTCTTA CCAGAGGTTT CCCATACTGA CAAGCTTGTA 43080
CTTGGCTGTT CAGACACTGT TTCCCTTCTC AGGCCAGAAC TGTTTAAAGC AAGCAAACAT 43140
GAAAGCCAGA AAAATGAGCT GATTGTGTGC TAACCACAGA CCCTTTGGTA CATGCATGTA 43200
CATGTTCCAG CATGCAGAAAT GACACAGGCA TTATACTGTT TTCTTCTGTG GCGTACACTA 43260
GAAAAAATG TATACAGTAA ACTCACTTTG TAAAACTTAC TTTGAAACCA TTATGTGCAG 43320
AGAGAAAAGC TACAGACCCT AAGTGTGTAT AGTTCAAGGC CATGGTCTCC AAGTCATTGT 43380
TCTATTGCTG TGCAGAGACA CCATGACTAA GGCAACTCTT ACAAAGAGC ATGTTACTGG 43440

GGACTTAATT AGTTTCAGAG GGCTAGTCCA TTATCATCAT GTCAGGGAAC ATGGCAGCAT 43500
GCAGGCAGGC ATGGCACAGA AGCAGTGGCT GAGAGCTACA TCTTGATCCA TGGGCAGCAG 43560
GCAGCGAGAG ATGGGGGAGG AGAGAGAGAG AGACAGAGAC AGAGAGACAG AGAAAAAGAA 43620
AAACAGAGAG AGAGATTAAT ATTGATTGAT TGATTGATTC TGGACCTGGT GTGGGCTTTT 43680
GAGATCTCAA AGTCCATCCT CAGAGACATG CTGACCTAAC TCACAAAGCC ACACCTCCTG 43740
ATCTTACCAA ACAGTTCATC AGCTGGGGAC TAAACATGCA AACATGTTTA TGGGGGCCAT 43800
TTTCAGTCAA CCCCCACCC ACAGCAGTAT TAGAAAATGA ACTTAGCTGA GTGGATCCCC 43860
TAAGCCTGTA GAATAGCACT TAGGAGGTAG AAGCAGGAGG ATCAAAAGTT AGGGTCATCC 43920
TTAGCTACAT ATTGAGTTTG AGACCAGCCT AGACTTCAGG AGATACTCTT TCTTTTTTTT 43980
TTTTTTTAAT TTATTTATTT ATTATATGTA AGTACACTGT AGCTGTCTTC AGACACTCCA 44040
GAAGAGGGCG TCAGATCTTG TTA CTGATGG TTGTGAGCCA CCATGTGGTT GCTGGGATTT 44100
GAACTCCGGA CCTTCGGAAG AGCAGTCGGG TGCTCTTACC CACTGAGCCA TCTCACCAGC 44160
CCGAGATACT CTTTCAAAAA GAAAAAAGA AAAAGAAAAT GAACCCAAAC AACTCAGGT 44220
CAGGAAATAG ACTATTAGAG CCCCTAAAC ACACACATAC TCCATCCATC CCCCATTCAG 44280
AACCTTCTTC ACATCTCCAA AAAAATGGAA CCATTCCACA AGTCTTAGTT TTTCTCTGAG 44340
TGTTACATTT GGGAGAATCC ATTGTTGTAT ATGATTGTGT CCCTTTGTTT TCATTGCTAC 44400
AGAATTTTCC TTTGAAAAGC TGAAGATATA GGACAGTGAT AGAGCACTTG CCTGGCATGC 44460
ACAAGGCCCC AAGTTGGGTC TCTAACAGAG CGATAAAATA AAATATTTTG AGAAACTACA 44520
GGAAATTTT AAGAAAATAC TTATATCAGT TCATTGAGAA TTTCATATAC TATATTTTGA 44580
TCATATTCAC CCCAGTTCC TCTTTCTAAC TTCCCCACCT CCCTACTTCC CCCATCTTCT 44640
TGTCATCATT GTTTTCTCCC CCCTCCCCC CTCCCCCTCC ACCTCCTCTT CCCCCTCCTC 44700
CTCATTCCCT TCCTTCCTCC TCCTCCTCCT CCTCTTTCAT AATGTATTGA CTCTAATTTG 44760
TTCTGTCCAT AACTTCTGG GTGCAAATTG ACTTACCAAG AGCTACACCC CTAAATACAA 44820
CTGATTTTCAT TTCTATCCCA GAAGCTCTCA ACTGTTTATA GGTCTCAGC TAAGGGTGAA 44880
GGCTCATAAA CTCTGCCCCA GTCCATGACA GAGTACTGCC TAGGCTTGAT CTTGTGCAGG 44940
TCTTATGCAG GTGAGATGGC TGCTGTGAGA CCGTGCGTGC ATGTCCCTGT CATGCCCAAG 45000
ATCCTGCTTC ACCCCTTGAA TTCTGGGTTC CCTGACCTCC AACTCTCTCT AAGATAGTAC 45060
CTGAGCTTTA GAGGTGGGCT TGATATGTAT GCCCCACTTG TGGCTGGGCA CTCCAGCGAT 45120
CACCGTCCAC TGCACACAAG AAGTTTCCCG ATGAGCTCTA AGAGCTGTAC TAACTTACGG 45180

ATACAAAGGC ACAGATTTAG AGGGCAGTTA GGCTGTGTCC TTTTAGC~~AAA~~ATAATAACAT 45240
 TGGCCAAATT TACAGAACCA GATATGTGCT GCCTCCGGTG GAATGGGCTT AAGTTCAGCC 45300
 AGTAAGTGAC TGGCTACCTC ATAACATTTG TGGCACTACT GCACCATGGG CATAGCTTAC 45360
 CACCCTGGTC ACTACTGCAG CTCACGGGGC TCACAGCTTC CTTTCTCTGA TATCCACACT 45420
 ATTGAGGACT ATTGAATATT ATTGAAGATT TTCCCCACAG CAGCCTGCAG AGTATCTTTG 45480
 AGTATGGTGA AGGTTAACA GCAGGGAGGA AGCTTCTTAG TACCAACTTG ATTTCTCCAT 45540
 GTCCTGTGAT GGGCATGTGT GGGTAAGCAA TAGGGTCTTA TCATCATGTT CTGGTAGGCA 45600
 ACCAAGCTAT GAAAGGCTTT TAGAGCTGGG TATAATGTAG TTCCAGCATT TAAGAAGTGG 45660
 ATCAAGAGTT TAAGGTCACC CTTGGCTACA TCATGAAATT GAAGCCATCT TGAGCTACTC 45720
 AAACCCTTGT CTCAAAAGCA AAACCTGATC ATCTATTCTG CATTAATCTA ATCAGCGTTC 45780
 TGATTGTTTC TGCGGTCAAG TTATTACAGA TAAATTTGTT TATGCTTTTG TGTGCACATG 45840
 CATATATTCT GCTTCAGTGT AGACCTAGGA GTAAACTGT TCATCCTACA CAATTGTATT 45900
 TAGCAAGTAG CAAGAGTTCA GGCCTTTTCT AACTTTCTGC CTGATTTTCC AGTTTTTCTC 45960
 CTCATTGTGT TTTTCTGCCT ATTCAGGATA TGAATCCTTT GTTGACTGTA TATATTGCAC 46020
 ATATCAGCCT AGAGTCAGAC AGTAATGACT AGAGAACAAA GCAACGCCTA AGGCACTGCA 46080
 GTTCTTTTCT GGAGGAATAG AAGTTAACAG CACCACTTTC TGGTTCCTGG TCTCTGGCCA 46140
 GCCAGGGAAT CCCTAAAGCT TTGATTCTGT TGATTGTCAC TTTGCTCTAA GATTATGACT 46200
 AAGGAATTGA GCTTCTAGAA TCAGTGACCA GAGTTCCTCA GATTTGGGAT AGCCACAGAT 46260
 AGAATCATCA ATGAAGTGT CTTTTTTTCT TTTCTTTTCT TTTCTCTTCT TTTCTCTTCT 46320
 TTTCTTTTCT TTTTTTTTTA ATCAAAAGTG TCTTTTAGGG ACCTAACTTT ATGGATGACT 46380
 CTTCAGCCCT TTCCACTCAT TCCCTGTGTG GTGTCATACC TCTCAGGGAA ACCAATCAGG 46440
 AGAGTTGAAT TCTGGACCCC ACTTAATCAT TACAAGAGAT AGTAAGGAAA TTCTTAATGC 46500
 ATATACCAAA TGAACATGCT AAAGAACTG GTGATTCTGC AGTTATGCAT GGATTCAGAA 46560
 ATCTGTAAGC CCCAGAGCC CAGAACATTT AATGTTTTGG AGTTCTGTGA TTGAATACTG 46620
 AGGATGCAAC CCCAAGATT ACAAAGGTCT CCCTAGAGGA GAACTGTAA CAAACCACAC 46680
 CAGTATGTTT GACATTTGCT CTTTCTCCA GTAGGCCCTT CCTCCAATGC CCTATGGTGC 46740
 TCTCATCTGC CCCATATGAT ATCTTCCTTT CTCTGATATC CATTGCCAAA ATGCTTTGTA 46800
 GCACATGGTG ACATGCTCTC ACCACGTGGG GAAGGGGTTA ATGGTAATCA GCATCTTTAC 46860
 TGTCTCTGAA TCTATAGTGG TATACACAGC TATACTGTTC TCTCAATTC CTGGCCTGAC 46920

CAAGTTGCTT CCTTTGCCTT CTCTGGGTAC CTGTGCCAGG CACACATCTG TCCGGGTAT 46980
 ACAGACACAC ATCTGTAACC CAGAGGTGCT CCAGAACCAA CCTCTACAAG CACATAGTCA 47040
 TCCGGTAGCC TTCAAACCCA AGGTGGCTTG TTCCTCTCTA AGACTTCAAG AAATCCTAGA 47100
 GAAGCTGTGA TCTTTGGGCC TGTACCCCAT TGAATGAATA GGCCACACAT TGCTGTCCAG 47160
 TAGACAGTGA GCCACAGCCT CTCTCTACCA GTATGCTGGA CCAGACACTA GGCACATTCA 47220
 CAAAGTGAGA GTGTCAAGTG TGTCTGCTCT AATCACCCAC CCCAGGCATC AGAGGCTTGT 47280
 GAACTCACA GGTTAGCCCT CCAGGAAGCA GGCCACAGGA CTTCAGGTTG AGCCTGGAGA 47340
 AAGGTGCCCC TGGCCGTCAC CTCCAGCAGC TACTTGGCAG GTAACCAGAA CATGCTTGGC 47400
 TCACTCAGCT CTTGGCTGTG CTCCCCAGAG GGAAGTGTTT CTAATCTGTC GTCCTGCTG 47460
 CTCCCATATA CTCTGAGGCA TTGTGGCTTT TTCTTGGTGG TTGGGCAGGA AGCCTCCAGA 47520
 GCCTAAAGGA ATTGCCATGC TTGATGACAG ACAAAGGCTA TTGATGGCTA TAAATCACTT 47580
 AGCTGCTGCC TGGCTTATTT AAGAGGAAGA GGACATGTTA ACTATTCTGA GGATAGGCCT 47640
 TCCTGTGGTG GGTACCCAAC TGAAAAGGGA TCTCACAGAT TGAATCCAGC TGTGCCCCGT 47700
 GAGTTAAGTG GAAGGAAATG CCCCACTTAG ACATGACTTT GCAAAGCCAA CCAGCAAATC 47760
 ATCCCATTGA CTTGTAGCTC CACCTCACTG GGCATCCTCA AGTGACCCAC CTTAAGCAGT 47820
 GTTGGGCCAG GATCCAAGGT GAGGAAGCCA GAGGCTGACT AGCTGGGACG GCACCACATT 47880
 GAGTGGGGGC TGTTCCTAAG GAGGCAGATC TGGCTTAGCC CTGAATGTGG AGACTGTGCT 47940
 ATCACCATCA TGTCCCTGAA GGCTGTCTAG AGCTCTCTGA TTCTGTAGTC ATGCCTCCCT 48000
 TGGGGGAAGT GCTCCACTCA CCGACCGGGG CTTTGTCTC CAAAGCTGAG ACATCTCCAT 48060
 CTATGTCCTT CTTGTTCCCT ATTTCTTCAC ATAAGACACT GTGACCACCT TCTCCTGGGT 48120
 GTGTGACCTA GCTTCGTTAG AGCTGTTTAG AATTCGAGAA ATACAATTGT CTTGTAGTTT 48180
 TCACTGGGAG AGGTCATAAC CTTTGCCCGT TAATGTATAT ATCCTCTTAA TGACATCAGC 48240
 TAGACAAAAC TAAGGTTTTA ATAAGTGAGG ATTGTTCAAA ATATTTATGT TATGTAAAAA 48300
 GTGTGTGGGT GTTTTTACAG TATGGAGATT GAACCTAAAA GTTCATACAT AGCAGGCAAG 48360
 TGCTCCACGA GCTGTATCCT TAGCTATTTT TAATTCCTTA TTTTGAGACA AAGCTTTTCT 48420
 AAATTTCCCA AGCTGGCCTA GTTATCCTTG ACCTTGGGAT CCTCCTGTCT TAGTCTCCAA 48480
 GTAAGATTAC ATGACTGCTG TGCCATGCCC AGCTGAAAAT GTTTTCTACT GAGTCTCCTA 48540
 CACTCTACAC AGCCATTTTC CCTACAGTGA GTGACCGCAG AGTCACAGGG TTTTCCCTTG 48600
 ACTTTACTGA AGCCTTGCCC TGTGTGTCTT TGTCTCTGCC CTGATGACTA TCAGAGCAGT 48660

TGTACCTCA CCACCTTCTA TGTGGTAACT GTGAACACTA GGCCTTGTGG GGACATAGAA 48780
 CCATAGGGAG AGAGGCAAAT GTTAGAATTC TCATCCCAGG TGAGAGAAGG TTATAGTTCT 48780
 GAGCCAAGAC TACCCTGGGT GCACCATACA GCAAAGTGCC TGTTTCATGCA GACATGACAT 48840
 GTTTCACACA GCTGCCTTTG AGGACACCTC CTAGTTCTGC ACCATCTTCC CCTCTCTGAG 48900
 ATTCTGTATG TTTGTGTTCT ACATCTGCCA ACTAAGCTAA ACTGACTCAA CTATTAGATG 48960
 CATTTTCCTA CCCCATCCCA TCCTATACCA CCCAACTGCA CCTCATTTCC CCCATCCCAC 49020
 CCCATCCCAT CCCACTCCTC CTCCCTCCCG CCAAATCCCA TCATGAAGTG CCTCCTTCCC 49080
 TGGAGCCTAG CAGGTTGCCC ACCACTTTAT GCTAAATATG TGTCTCTAT CCTTTAGTAT 49140
 AACCAGACTA GTCAGGTGGT CACCATGTTT TGTGTAAGGA ATGCCATTCA TCACTGTTCT 49200
 GCTCATGAAA CAGAAATGCCC TTTTCACTCC CTCTGACTTT CTCAGTGAAT TTTCCAGTGC 49260
 TGATGTCATC AAAGTTGACT CCCAATTTTT AACAAACCTC AGTCTCAGAA CTACCAGTCC 49320
 CCTGCTGAGT ACTTCAAGAG GCGGGTCTTG CCTCTGCCTG TGCAACTCAG TGGAATGTGA 49380
 ATGCTTTTGA CTGTGAGGTA GAGAGTGCAT ATTAAGAGGC TTTGCAGATT TTCTGTAGAT 49440
 TCTGGTTCCC AGTACTTAGA GCAGACCTGG GACCCAGCCA GGGGCTGCTG AGGAGTTTGT 49500
 AGCACTGATG AAGTTCTGAA CAGTCCCTCC AGCAGAGCTA GCACACTGCG GATGCTCAGC 49560
 AGACACCGGG TGCACGCCTC TCCTCGCAAG CATGGATTGC TTCCCCTGCA TCCTTAATCT 49620
 TAGCATGATG CCTCCGTTTC TTCTAAAGCA CCAGGCGCCC GTCTCCTTCA CTTACTCTAG 49680
 ATGGTTCTCA TGGTGGAGGT TAAGAATTCC CCATCTGAAC TCTAAACCAA ATACCTTATG 49740
 AACTTCCAAG TTTTAGATTT TAGAGCATTT GAGATTTTAT GTTTGTATTC CAGAGCCTAT 49800
 GCAAATATTC ACAAATCTGA AAATGAAATC TGAAGCACTT TTGGTCTCAG CATTTTCAGAT 49860
 AAGAGGTTAA CAGCCTGTAT GCTAATCATA TTTATGGAAT ACTTAGCAGT GTGTTGGCCC 49920
 CTAAGATAAG AACTGATGAA ACATCTACAC CTCCTGGAA TAACCTGAGA TTCCACAGAC 49980
 CCTGTGGTGT TTGGAGCCC 49999

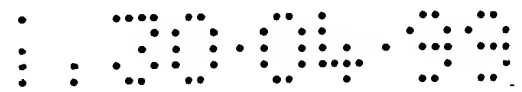
(2) ANGABEN ZU SEQ ID NO: 6:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 36901 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(iii) HYPOTHETISCH: NEIN



(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

(A) ORGANISMUS: *Mus musculus*

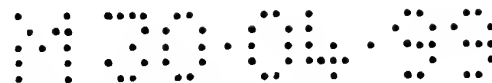
(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

CATTCTGTG CCCATTGAGT TACCAAGACC AGAAAACCAC TATTGCCATT GGGCTCTTGG	60
GAAATAAAGG TTCCATTAC ATAAGGATGC CCACTCCACA CCTACCACCA TCATTTTTCG	120
AGTCCCTTCC TGTTCAAGCA AGCTCACCAT GGGAGCCAAG CCAGTGCTGT TCAGATCCCA	180
GTAGCAATAT CCACAGCCAG AGAGATGCAG AAGTCATATA GGCAAGAGCC TATATGCGGA	240
CTGTTACATA CCAGACAGTT GTGTCCCCAC TGCTAAACCT AGAGAAATGT TCCACAAATG	300
GCCCAGATTG CAAGAAGAAC CCTGGGAAAT TCTACCATGC ATCTCACAAA TTAGAAGACC	360
AGTCATTGTG TGTATTGTAA GATCAATGTA AACCTCATGC CTTTGCTTGT CTAGCTAGAG	420
CCAAGCACTG TGCAGTGCAT GGAAACAATA AAGGTCCAGA GAACCCACTG AGGGAGACAG	480
GCATGGAAAG CAATATTTAT AACAAATACT TAGGGTGGGG CATGATGGGA GAAATGTCCT	540
TGGGCTCAAT CAGCTCATGA TCAGATGAGC GGTGTGGTGG AAACACGAGG TGGGAGCAGC	600
ACAGGTCACC CAGCTGTGGC CAGAAAGCAG CAAATGGCAA GAGGAAGGGG CCAGGAACAA	660
GGTATAGACC CCAAGAATTC CCAGAACTCA GGCCCTGAAG TGCCCTTCC TCCTAAATAC	720
TCTGCCATCC TCCAAAACAG TGTCATCAGC AAGGGACCAG GCCTTTAACT CATGAACCTC	780
GGGGGGGTGG GGGGGGCGGC ATTTCATGTT CACACCATAG GGGTGACAAA GGAGTTAGGA	840
GCCAGGCTCC CAGGATGCCC AGCCTGGGAA GGAAAGTACA TGCACTGCTT CTCTCAGCTG	900
GGGCCTCATT GGACAGGCAA GTGCCCTGTG AGCAGGTGTC AGGTAGGAGC CTGTATTTTG	960
ACATGGAGAG GACAAGGCAG GTGCCTGGGT GCTGCCAGGT GGAAAGGGCA AACGGCCTGT	1020
GTGTGTGTCT GGTGCAGTCC AGGCACGTGC AGGGGAAGCC CAGAACTCGC TGGATGGGAA	1080
CACACCCATC TAAAGCACTC TGAACCCAGT TCATAAAACC ATGGGTCAAT ATTTTCAAAG	1140
TCACAGAACT AATGAGCTCT GCCAGACTCA ACAGACCGCA TCCCAGTGGG TGATAAGACA	1200
AGTGTTAGCA CAGAGGAAAC GGCCCAGGCG GGAAGAGGCT TTTCTTAATC TGTTGGGTTT	1260
CGTGTTTATA GTAAAGCAGC TGCCCTTGGA CAAGAGTATT CATTTATCAG GTCACCCACA	1320
AAGGAGGCTT AGTTACTATG CTCACCCTGT TTGGGTTTAA GTAATAACTG TCTACAGACA	1380
AGTAAAAATT GGATCAGGGC AAGTTCAGTA GGTCCCATCA GGCTGCAGA AGCTGTCTCA	1440
GGCTCTGACT GCCAAGTTCG TGTGCCTGTT GTCCAGCAGG AATAGGCAGA GAGAAAGCTG	1500

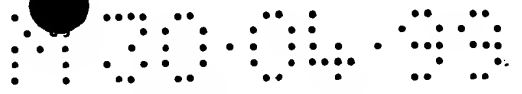


TGAAACCCT	AGCCTAGCCC	CGAAGAGCTC	TATTTTCACC	CTTTAAAAAT	GTGTGTTGTC	1560
TTCCACTCAG	TATTTCTGTG	AAACAGCAGC	AAAGAATGAT	TCTAGTGTGC	TCATTTAGTC	1620
CCTGAACAGT	TCATCAGCAT	CCCACTTGTC	TCTGGGATTG	CCAAGACCAT	TCAGGCCTAG	1680
ATTCCCCCA	CACCTTCCTT	CCCACGGCTT	GGGGTCTGCA	GAGGAAAGTG	GGCAGAGGAA	1740
GGGAAGAGC	CAGCTCACAT	TGGTAAGGCC	TTACCAACCA	GGAAAAATAA	GGATGGCAGT	1800
GACCCAGCTA	AGCATCCTGA	GTACTACAGA	GGAGGCTTTG	TGAGGGAGGC	CTCACTTCCA	1860
ACAGAGATTC	TGTCACCTCC	TGAGTCCTGG	ACTAAGGTAC	CCAGAGTCAC	CTTCTCACTC	1920
CCGCTAGCTT	CTGTGGGTTC	AGTGACACAG	ATCAGGACCC	AGGCTGTACC	TGGAAGCGTC	1980
AGTCTCACGA	GAGGTCTTAT	CTTACTCATT	CTCTGTTGTC	TTGAGGTAAA	AACAGCATGT	2040
GCAGAACTGT	AAGGTGCTGC	TGGTCTTTGT	AAATAAAGAA	ATAATCTCTG	ATGAAAAGTA	2100
TTTAAAGCAT	GGAAGTGCAC	ACCTATAATA	CCCACACTCG	GGAGGCACAAA	ACAGAAACAT	2160
TGCCATAGGC	TTGAAGCTCA	CCTGAGCTAT	GTAGTGTAGC	AAGTTCAGA	AGATCTGGAC	2220
TGTATGGTTA	AGACTGTCAC	CACCATCATC	ATCATAATGA	ATTGTATATT	ATTATAATAA	2280
TATTAAAAAG	TATTTAGTGG	CTGCTTCCTA	TGTCCTAGTC	ACTGTTCAAG	GGACTGGGAG	2340
GTAAGCTGTC	TGAGCTCCCC	AGGTTAGTGA	CATTGAGCAG	CTGTGACTGG	CCCAAAAGAA	2400
TGCAGGGACA	GGAAGAACAG	GAAAAAATC	ACAAGTAGTC	AGGTAGAGCC	CCAAGCTAGG	2460
ACTGCAGTAG	GCAGAGCAGG	AGTGAGCAAG	CTCACACGGG	CACCACTAAG	AGCTGATCCA	2520
ACCATGGTTT	GTCCGTGACT	GATGGCTTTG	GAGCAAAGCA	AGGATACAAG	TAGAAGCCAC	2580
ACTCCAACCT	AAGAGTGTCT	GGCTCCAGGA	TGCCCTTCTC	CTGAACCTTG	GACTTCTGGT	2640
GAAAACTTAT	GGATGGTGGA	TCCCTAATGG	TTTCCCAAGT	GCTTGTCTTT	CTAGGAAGCT	2700
TATTTTAAAC	TCCACCCCCA	TGCAAGGTCA	GGCTATGGCT	TACTCAGATA	CAATCGTAAA	2760
TGTCAGCAAA	GCCATGGAGA	AGATGAAGAA	GTAAGAAGGA	TCATCTCCCT	TTTACCCTCC	2820
AAAGACTGAA	GCCTGTGGAC	AGGGCCCTGG	GCAGTTCACC	CAGGGGCTTG	ACAACTTACA	2880
CAGCTCTGAC	TACGTTCCCTA	TGCCAGATGC	AGTCTGTCTG	CTCCTCCCAT	CTGTTCTGGT	2940
CTTCCCCAGA	GCCTCAGACC	AGCAGACAGA	AATCAAGCCA	TGCTTGGTTC	TAGATCTGTT	3000
GCAGGTGCAG	TGTGCATGGT	GGGAAGGGGA	ATGAGGCAGA	GCAAGCAGCT	TGAGTCACTC	3060
ATGCCAGGGC	TCCCTCCACT	AATATCCCTC	CCTAGAGATG	GACTCAGGTT	CCTTCCACAG	3120
CCTCTGCAGG	CCTGGTCTTG	TATTGCCAG	ACAGAGATCA	CCTACTTCAG	AAGGGGCACT	3180
CAGTACTTGC	AGTGTCTCT	TGATTGGATG	GAACCAAACA	ATGCTGGGAC	ACAGGCCATC	3240

CCCCAGACCC ACAGGAGCAG CTCCACCATG CAAATCTACC TCCAGCTTGA GGTGGGCTGC 3300
 ATAGGTAAGC TGATACACAA CCCTGCTTGG TAAAGGAGAA GACAAAGTAA CATTCAATAC 3360
 AAAAAAAAAA AAAAAAAAAA AAAAAAGAG TTTGAGGGTC TAGACCAACT AAGGCTTGGA 3420
 GTTCTTTAGG GAGCAGCATT TGGATTTTAT GTACCATCCC AGAGCAGGGT TCTCCAAAGA 3480
 GAATAGCTTA TACCTCCTTC CCACTTAACA CAGCCACCCA AGGCCAGAAA ACCTAGAGAA 3540
 GCCAAAGCTG CAGGACTTGG TGGTGCCCCA CCCAGATCTG GGCCCTGCCA CATTCTGGCT 3600
 CTAGTCGTCT TCTATAGCCT CTGAGACTCA GTTTCCCACT GTGCACATTA AGACCTACAG 3660
 TTTTTTTCCT GGGAAAGGAC TCATTGGGCT AAATGACAAA GCACACAGAG AGCTTGGCTG 3720
 CACTCTCTTT TCTTCCCACC ATTAGTGGCC TCACCACTCC AGGGTGGCCT TGGAAAATGG 3780
 GGCCACCCCC GCCCCCCCAG CAGCCCAAGC AAAGCACACT TTGAATAAAG CAGAGCAGCC 3840
 TGAGCTCCCG GGTGACCTGG CTCCTCCTCT CCTCTCTCCT CTAGAGCTAT CTCTGCAGT 3900
 TGTATGTGTA TGAGAGGATC CGTGTGTTTA AAACACCCTT CTCCCTAGAA CATCTTCATA 3960
 CCCAAATTCT AGCTTTCAAA CTAAAGTTGA TCCCTCCCAA AGTGAGAGGT GACTTTGGCT 4020
 TCCCTGAGTT TATCCAAGCT CTGTTCTTGG TATAGGTCTT CAGGGTCAGC CTCCTCTACT 4080
 TGGGTGTAAG AGGGAGCCCT GGCCTTGGCT AGGATCTGAG CAGGGCCAGA AAGCTGTTGC 4140
 AGGCAGGCAG CAGCTCCCAG AGGGAATGTG CTTCTGTGTG CCTTGGCCAC ACCTCCTCTA 4200
 ACCAGTGGTT CCAGTTTCAG TGGAAGTAGA GAAAGGCTCT CATGTGTGTG TGTGTGTGTG 4260
 TGTGTACACA TCATAAAAGA GCCAGCAAGG CCCAATTACC CTTCACTGCA ATGCTACACA 4320
 GCACAATGCC TGGTTCTGCT TAGGGGCCAG AGCTGTTGCC CACGTGCAGG CCTGCCCCGT 4380
 GCCTCTGTGT GCAGAGCTAA GCCTTGGGAA GAGCAAGGCT TCGTGGCTAG CTTTATGCTG 4440
 ACAAAGGGCT TTCAGTGCTG TCAAATGACT GCAAGCAGTC CCTTCCCCCT CCCTACCACA 4500
 GCCACTGGGC CTCCCTTTGG CAGGGCCAGA GGGCTGCACT TGAACGCCTA GCCTCTGGAG 4560
 ACTTCCTTTT GAACTAGAAA AACATGGCTC AAACATGCTT CACTGCAGCA GGGCTCTGCC 4620
 TGCTGAACCT ATAGAAAGGC CTGGAGTAGA TTCAGTCCCA CAGACTAGAA AACCTGGCTC 4680
 TGGCCTCACC CACAAGGCCT GTTATGTCTG GCTCCAGAGG CCTGCTCCTC TGGGGTTTTT 4740
 CATGCCTGTG AACTAGGCCC CATTCAATTC CTGCGGTTT CATGGGAACG TCCAAAATAT 4800
 TGAGCAGGTT GCAGGGAGCC CAGGAGGAAA GGGGTCAGTG AAAGGCCCTA GCTGTGACGT 4860
 GGGGTGGCCC TGTGGTCAAG CCCTGGTGGG CGCCTTGTCA GTCTGCTGCT GCCTCTCCTC 4920
 CCAGGCACCC CTTCCACTCC CTTGAAGCTT GGCCTGCAGC AGCACTCCCC TTCCCCACCC 4980



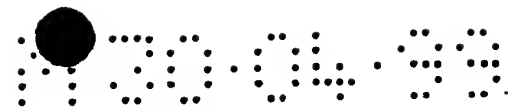
CCAGGCCTCT	ACTTTCCAGC	TCCCTAGCCA	CCAGCCCCAC	CCTGGCCTGG	CCTCAGAGGG	5040
AACTGCAACA	AGATCTCTAC	AGTTCCCCAC	CCCCAGCATC	CCTCAATTTA	GTACTGATCA	5100
GACCACTGAC	TTCCCATCAC	GCCCCATTCC	CTTGCACTTT	TCCACCACAC	TACACTCAAT	5160
TTGGGGCTGC	TGAGAGAGCA	GCAGGTCTCC	TGTGAGGGTG	GCTGCTGTCT	TCCCACCTTG	5220
GGCTGCCCAG	CTATAGAGGA	GAGTCATGCT	CTAGCACACA	ACTCCTGTGA	GAGCCCAGCA	5280
GCTGCCTTCA	CAGCTACTGG	GGAGCCCAAG	GGCTCCTTAA	GCCAACAGTG	AGGATGTACC	5340
CATGTGGGGG	AAATTTGGTT	TGCCGAAGAA	ATGAATTTGA	AACTAGCTGG	GAGCAATTCT	5400
TATCAAATTT	CCATGTTAGC	AGTTTTCACC	AAGAACTAAT	TGAACAATCT	CTGTGAGTGG	5460
CCTAATTCCA	TTAGCATGAG	ATTCCCACAA	AGTTAACAAG	TGCCCTAGTG	GCCAAGGGCA	5520
GAGAGGCTCT	TCTGTCTCAC	ACTTGGTTTT	GGTCTTTGAA	GATGGATGGA	GTTTCAGGTT	5580
TCAGCAACAG	CCAGGCAGAT	GCTCACCTCT	GGCCCAGTAG	GCTTCAATCT	CAGCAGCTCA	5640
GCTCCAGATC	AACTTCAGAA	GCCACTTTGC	AAGTATTGAG	GGTATGAAAG	GGCTGATCAG	5700
ACCACTGACT	TCCCATCCCA	AGATGAATTT	CTCTTCTGGG	TTAGCAGGTA	AAATGGATCT	5760
GAGGGTAGAA	CATCCTACAG	ACCTCACCTC	CCTTGCCAGG	CAGTATTGAG	AGACCAGGTA	5820
CAGAGGAGTA	GAAAATATGA	AGGCAAAGTC	TGAGGAGCAT	GAGTCTGGAC	AGGGCCTGCC	5880
CTCAGCACCA	CCTCCCCACC	TGAGGCAAGA	CCCAAAGTTA	GTGCCAGCAT	CTCACTGTTG	5940
TCCAGAAACT	GAGTTCTAGG	GGCAGAAACA	GCAGCCACCT	GGGACCTGTT	CCTGTCCTTG	6000
AGCCACAGCG	AGGTAGCTGT	TCCCTAGTGG	TATAGTACTT	TCTCTTCTCT	CCACTGCCCA	6060
GTGGGCTTGA	CAGTTCCAGG	GACGGTGCTC	TGGGGTTACC	CATCAGCCCT	GTGGCATCAT	6120
GCTAGATGAG	GAGCCCAGAG	AATGAAGCAT	CTAGCTTCTT	TGTCCCTGAC	TAGCTATAGA	6180
CTGAGCAAGG	GTCCTCTCTT	CTTGACAGCT	GCAGCATGGT	GTCAGCATTG	ACTGCTATGA	6240
ACCAGCCTTC	CTATAGGTAG	CATGGTCAGG	ACAGAGGTTG	CAGACCTACC	TACAAGGCCC	6300
TTCTTAACC	TGCTCTACAA	TGAGACATAA	GCCAGTGACT	CTTCCCTTCC	CCTCCTCTGG	6360
GCCTGCTGGA	TGGCTTCCTG	CGGGCTCTCT	CAGGGCATGA	GCCCTTGCCT	CCTAGAATAC	6420
CTTCGACTTG	TCTAAACTA	GTCATAAGGC	CCTGGCTCCT	TCCTTCTGTC	ACTGACTCAC	6480
CAAACTCAA	TGGAGCATTG	CCTGCACTTG	ACCTATCACC	CCTTCCCTGT	TTTCTAAAC	6540
CAGATTCCCC	AGCCCTACCA	CCCTGGTGGT	TTGCCTCAAC	TTGCCAGCCT	CAGGGGCCTT	6600
TTCTTACCCT	TTCTCTGCC	TCTGCAGCAC	TTCTCACAGG	GCAGCCTGCT	ACAGCTCCTC	6660
CATGTCCCTC	TGCCTTATTC	TACCACCTCT	ACCTTCTCTG	TTCTGGCCTC	CTGGGGGCCA	6720



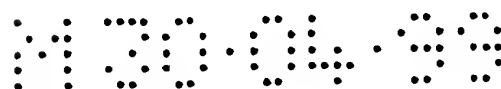
GTGCACACGC	CTTCGTCACC	TGGCTCGCTC	AAGCCCTCCC	TTAATTGTCT	CATCCCTCAT	6780
CCGGTCCTAC	TCTGTCCCCC	AGCCCCAACT	ATTCCCACAT	ACTTATTTGA	AACATCTTTC	6840
TTGCTCAGTA	GCCTTCCAGC	TCCTGAGTGG	GGTCCAAGCC	TGTACCCTCA	ATTCCTTGCC	6900
TTTCCACCTC	GAGCTTTGTG	TTTCATTTCT	GGTTCCTTGA	CATCCCTTGA	AATGAATCCT	6960
GCTTGTGAGT	GTACCTCCCT	GTGGATGGAT	ATACCTGTGG	GCGTCTTAGG	AAGTATTTAG	7020
GCATTCTGAT	TGCCTCTGAG	GCCACTGGCC	CCAAGAGCAC	AGACTGATGC	GTAGGGATAT	7080
AGGACTTGGA	GCAGATCACT	TCCCTATTTG	CACATTAAGC	TCCTGCCACC	CAGAAAGATA	7140
AGAACATTGT	AGGGCCATAG	GAGAAGTGAT	ACCCAGGGTG	GAGTGAGGCC	ACAGCTAGAA	7200
AAGATGAGTA	AGAAATCCAA	CAAAGGGATT	CAAAGCTAGC	TCTGAAAGCT	GAGGCCTACC	7260
AGCCATTGCT	AGTGTAATA	ACTCTGCTGC	TGTGTATGAA	GGAAGTAGTA	CTCAGTAGAT	7320
AAGGAAGTAG	TACTCAGGAG	ATAAGGAAGT	AGTACTCAGT	AGATTGGTTA	GGGCCTGTAG	7380
AGAAAAGATC	AGGAGACTTG	GTGACCCCAA	ATTATCAGCA	TGCCTGGCAG	TGAGTATTAG	7440
GAAGTTAGAA	ACACCTGAGA	ACTAAACAGA	AAGGACAATA	GTGATAGAGG	GACCCAACAG	7500
TCCTACCTCC	TGAACTGGAG	CCTGATGCCA	TTGCTCCCAG	GAGTCCTTCA	CTCTGTGCAG	7560
GTTGTTGAAC	ATCCACTCTG	GGACTAGCAC	ATATACCACT	AGGGATGGAG	ACGAGATACA	7620
ACCTAGGACC	GAGAGAGGCC	ATCACAGTCA	TGAAGGCCAG	ATGCTATGAT	GGGGACCAAG	7680
AGGATGCTAA	GAGAGAGTTC	CTCATGCTAT	CTTCCAAACT	GAGTGATAGC	CAAAGAAAAGG	7740
ACATGAGCGA	GGAGCAGCCC	TAGTACTCTG	GGCTGTGAGA	ACAGTATATG	AAAGGACAGA	7800
AGCCAAAAGG	GCCTCAGGAC	TTCAGTAGAG	CCAAAGTAGG	ATGGAGCAGG	GAAGAAGAGT	7860
GATGCAGTCC	AAACATACAT	AAAACATACC	ATATTGTTTA	GCCAGGTAGA	GGAAGTCTTA	7920
GTCTTAAACA	GTGGTTCCTG	CTGGAAGGGA	CATGACCCTG	TTTTGTGTGA	AGGCAACACA	7980
GTAGCAGGAG	ATGACGACCT	GGACAACAGT	GATGACAGGA	AGGAAAGCAA	GAGATGCTTC	8040
TGGAAATCTA	CTCCAGATCC	TAGAACTGGA	CCATTTGAGC	AACTCTTGCA	TACCCTGTTG	8100
CTCTTTAAAA	AGAGGAAGAA	AGAAAAGAAA	AAAGGAAAGG	AAAGGAAAGG	AAAGGAAAGG	8160
AAAGGAAAGG	AAAGGAAAGG	AAAGGAAAGG	AAAGGAAAGG	AAAGGAAAGG	AAAGGAAAGG	8220
AAGAAAGGAA	GAAAGGAAGA	AAGGAAGAAA	TGAAAGGGGA	AGGAGGGGAG	GGGAAGGGAG	8280
GGGAGGGGAG	GGGAAGGGAA	GGGAAGAGAA	GAGAAGAGAA	AAGGAGAAGA	AGAGAAAGAG	8340
AAGAAGAGGA	GAGAAGAGGA	GAGGAAAGGA	AAGAAAAAAA	GCAATAACAG	GACAGGTGCC	8400
CAGACAAGAG	GAGGTCTAGC	TAGGCTAGGG	TAGACACACT	GTAGTCTGAG	TGGTACTTAT	8460



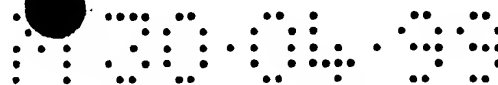
TTATGGCCAG	GAACTTGGTC	GCTGATTTTC	ACTTGGTTGG	CATGCCTGCC	TTCCTCAGAG	8520
GCTTCTCACC	TAACCACTGT	CTGACCTGTC	AGGATGCTGA	GGTTATGTAG	ACTGAAAGAC	8580
CCTACATAGA	GAAAGACACA	ATCTCAAAAA	ATTAGGTAAA	TAGCAAATAA	TAACCACATT	8640
TGGACACAAG	TAAATAAACA	TGGCCCAGTC	TGGGTCCTCG	GATGGTAGGT	GCAGTGTCCA	8700
GCAGCATAAG	TTGTGTTGAG	CATACTCACT	TCCTAAGGTA	AAGAATGCCT	ATAATAGTAA	8760
TAAATTGACA	GCAGTGTAAG	TTTGTATCTG	AACCTTTCCC	TTTAAGTGGT	ATCAGTACCG	8820
TTCTGGGCGG	AAGCTTCCTT	TCTTATGACA	TGGAATGTGC	ATCTCTGGTG	TGCACTTATA	8880
TATAGGTTGA	TTATGGCTTG	CCAGGACATG	AAACCCTGGC	TCAGCTGGTC	CCTGGGATGA	8940
GAAACAGCAA	ACCTTCCCCC	TCTTTCCCCA	GGCCTTGACG	GCCCAGACAG	CAGGTAGGGA	9000
CTGCTTGAGA	GAGGGCTGCA	GAGCTTTCAC	CGTGATGTCC	TGGCTGACAG	CCTCCTGTCA	9060
CAGAAGAGTC	CTACCCAAGA	CCTCCAGAGT	TGTGGGGCCC	CAGTGGCTCA	GGCCTCCAGA	9120
TGCTCAGCAG	ATGCCAGACC	TGGGACTGAG	GCCCCATCTC	TGAGGGCTTG	GCTTGCTGTT	9180
CTGGAAGGTG	ATCCTGGCTG	TCAGCCATTC	TTGAGCCCCT	ATTTAGAGCA	GTTGTCAGGC	9240
AGTTGCTGGG	ATTCAGCTAG	CTCCCCATCC	CCAGCAGGGC	TGAGTGATCT	CATGCCTATG	9300
CGATGCTGTC	GCCTGGGGAG	GAGGTGCCCT	AAGACTGAAG	GCAGGTGCCC	AGACCAGAAG	9360
GAGAGTCTAG	GCCATGGCAA	CCCAGACAAC	CCTCAGCCAC	TTTCCCAGTT	CCATACCCTA	9420
ATGTGCTCCA	GCCTGGTTCA	TTTGCCCTGG	GATAGCACAA	GGCATCATTT	GAGTTTGGCT	9480
GCAAACTTTA	TGTGAAGTTT	GCCCCTTTCC	CCACAAGAGA	GGAAAGCTCA	GATTGATAAG	9540
CTCGCTTGCC	AGAGACCCCA	CAGCCAACCG	GTTTGCACAG	AACCCTCAGC	CCAAAAGGCA	9600
GCTTTAGCTA	ACGAAACAGC	AACCTGGCACT	CCAGGGACCC	CTGGACTTTG	GGCCACAATT	9660
TGTAAACTCT	CGAGCTATTC	TTCCCAGAAA	GTTCTTGGGT	TCTAAGTGGC	TTTTGCCACG	9720
TCCCAGGACT	GGAACAGAAG	AGTCTGGTGG	CCCCCTGCTG	ATCACTGTGA	GAAGTGCACA	9780
AGGGTAGACA	GGTGCCAGCA	AGAGGGGCCT	TGGCTAGCCC	CAGGTGAGAG	GAGAGATCTG	9840
TGCACCCCTC	CATGGGTGAT	TGGCCCCACA	GGGAATCTTA	AGTTCAGTGG	AGCTCTGGCT	9900
GCTGCTGGTT	TGGCCATGTC	TCAGCCTGTC	AGTTCTAGAT	CTTCTAGATC	CTGGGCCTCC	9960
TGGGAGTCTG	GGAGCTCCTG	GGCCAGAGTA	TCGCTGGGTC	CTTTGTGATG	TGCACATGCT	10020
TGCTCCTTCC	CCTTCCACTT	GCAGGATGAG	AGGATTTTAA	GATCATTTCC	TCAAACCACC	10080
CTAGGACACT	AACGAGCCTT	ATCCGCACCC	AGAAGTGGGA	ACTTTGTTCC	GTGCATCCTC	10140
TTGGTTGGTG	ACAGGATTTA	AGTTAATGCT	TTGCTCTTGA	CAGACTGTTG	TGAAGAATTC	10200



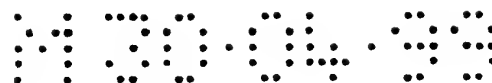
CTAGGCTGAT	GTCTTAACTC	AGAGGGAGAG	AGGAAGCGAA	GGGCAGATGG	ACAGGGGGTG	10260
CAGAATGGAC	AGATGGACAA	GGGCTACTAA	TGGAAATAGG	AATCACAGGC	ACCAAGGTGC	10320
CTGAACAAGG	CCAGCCTATG	CAACCAGAGT	CATGCCAGAT	TGTGATCAGA	GTTAGACATG	10380
CTCTTCTTTT	CTCAAGGTCT	TGGGCAGCTT	ACAGGGCTGT	GCAGATGTCC	ATGGAGGATA	10440
AATTGTCAGG	TCATGGTCAC	TGGAGAAGCT	GCTTGCCTGG	AGTCTTCTCA	TGCCTGTTTC	10500
CCATAGTGGC	CCCTCCTTCA	CCCCATCTCT	CTTCTCCCAC	CATGAACTCA	TGTGGAACAA	10560
AGCAGAAGAG	TTCCTGTGGA	CCAGGACTCT	GGATCATCCC	ATCAAAGTCT	CTGACTTATA	10620
GCTTGGAGCA	TGGAGAAGGG	TCCCTGTCCCT	GAGCCATTAG	CCCACCCTGC	TCCTGCCTGC	10680
CTAACAGCCT	TATCCTCACA	GTCCTGCTGT	GGGGCCCTAC	TGCCACCTGC	CGGCTTCATT	10740
TACAAACTGC	AGTCCTAGTT	CAGCCTTGGG	ATTACAAGAG	ACTGTGTACT	CTGGTCAACA	10800
GGATTCTGAG	ACTGCACAAA	GAGAACAGGT	CTGGAAACAG	TCCTGACTTC	CCATAGCAGT	10860
GTCAGAGCAT	TTATTTAACA	GTCTGAGCAG	GGACAGACAG	CATCCCAGCA	CTGTGGAGGT	10920
TGTGACAAGG	TGAAGGATTA	TCAGATGTGT	TAGTCATTTG	TGTGGTGTAT	GTGAAGAAAG	10980
GAAAGCACCA	CTGTGTCTTG	GACAGTTGAT	ATTCCTGCTT	GGTATCTGGC	CCAGAACACA	11040
TGTTCCCTCT	GCCTTTGCAC	CAGCCCTGTG	ATCAGACATT	AGCATTGTCT	TACTTTGGGA	11100
AGGAAGAACA	GGAGATTACAC	CAGGGGTTCC	ACAACAAGAG	TGTGGTAGAA	CCAGCATTCA	11160
AACTGTCTCA	GAGGCTTGGT	GGTCAGTGAT	GGTGATTGTC	AGTACTGATA	AGCACAAGAA	11220
GGGATTGGGG	ACTGAGATAA	GGGTGTCAGC	CTAAAAAGCT	CTGCCTACAA	ACTAGTGGGT	11280
AACACAAAGG	CTTTTCTTCT	TGAGCTGAGT	CTAGTGAGTC	CATGACAGAA	GCCAAGTGTG	11340
CAGAGGCCCC	CATGACTGGA	GCTAGGCTTG	CCCAGGCCCC	AATGACAGGA	TCGGGTGTGC	11400
ACAGGTCCCC	ATGACAGGAG	CCAGGTGTGT	CCAGACCCCA	CCTAGTGGGC	TTCATGAGCC	11460
CCTTGTAGAG	AAAGCTCTGC	AAATAGGCAC	CTAGACAGAG	CAGAGGCAAG	CGTCTTCACA	11520
GCAGGTCCAG	TCTGGAGAAG	GAACATTCTC	CTATATGTCT	GATTTTCCTT	CTAAGAACTT	11580
GTCTAGATGA	CAGATCTGAC	CAAGCAACAC	TACTCAGCCT	CCAGTAGAGG	GATTTATCCC	11640
AGGTTTCTCT	AGACACTGGC	AGACTCTCAG	AGCTGCCTCA	GTGGGAGAAG	AAGACTAAGG	11700
CTCAACATGC	AGCTTGGGGT	GTCTCCTCGA	AGCTGAACAA	GGTCTCTAAT	GGCTTTTGCC	11760
TTCCAGGGA	GCAAGCTTTT	TCCACACAGG	ACATGCTGAC	TATAGTAGTA	TCAGGATGTA	11820
CACACCTGAA	AGACTTCATG	TTCAATCCAC	TTATTACCA	AGGGAGCCCC	AAGGGTCAGG	11880
GGAGAACCTG	CCTGCCCAGG	ATTGAAATAC	AGGTAACATA	CTTCAGGGCT	GGTTGACTCT	11940



GTCTCCTGCT	GTGCCTGGCT	TCCTACCCCT	GACACACTTC	CTCCATCTTC	CATCAGTCCC	12000
CACCTCTTCT	CACTAGGGCC	TTGACATATT	TTCATCTTCC	TATTTAGAGC	TTTATCCCCA	12060
TGTACTTAGT	TACTTATAGT	AATTCTAATT	ACACTGAAGT	GAAGGAAAAT	AGAATGATAG	12120
CTCTTCTTAC	AAGTGAGCCC	CAGAGGAAGC	CCAGCAGGTC	TTCTTACCAG	AGATCATTAC	12180
TGTGTATCAT	CTCTGGACCA	GGCATGACCT	GAGAGCATCC	CCATTTAGTG	AGAAATGAGA	12240
CAGGAGACCA	CATACACATT	CAGACCAAAA	GAGAAAGTCA	TTATTGACAG	GTTGACTCTA	12300
GGAAATCTGA	GCATGGAGAT	GAAAGAGAAA	GAGCAGAAGA	ACTAGTTTGA	TCAGGTCACA	12360
GAAAGGTTCT	TACACTGAGA	ACTAAGGTAT	TAGAGAATCA	GCTGAGCCAA	GGCCTTGGGA	12420
CAGGGGCAGT	AGCACCTGTC	TCCAGGATCC	CTCTAGTTAC	TGTCTATCCT	CCACAGGCTT	12480
GTAGAGGAGT	TCATGCTCCT	GGCCAACATG	GCGGTGGCCC	ACAAGATCTT	CCGCACCTTC	12540
CCTGAGCAGG	CCCTGCTGCG	CCGGCATCCC	CCACCACAGA	CGAAGATGCT	CAGTGACCTG	12600
GTGGAGTTCT	GTGACCAGAT	GGGGCTGCCC	ATGGATGTCA	GCTCTGCAGG	GGCCCTAAAT	12660
GTGAGTGCTA	GTGGGCAGGT	AATGGGAAGA	CCTGCTTGGA	GAAAAGAGAT	TAAAGCCTAG	12720
AAGTTGGGCT	GGTGGTGA CT	TGTCTGCCTC	CATGTAGCCA	CTCCCTATGT	AGCCAGGTCA	12780
GTCTCCCCTG	CGGTGGAGAA	GATGGCATCC	ACTAGGGGTA	GGCTCTATTA	TCAGGTCTGT	12840
ACCAAGGGAG	ACTATTCAAG	GTGTAGCCAC	TTGCATGGCC	TCTAGCAAGG	ACTGGACTGG	12900
TCCTTGCTGA	GCCAGGGTAA	CAGGAAGCAA	GGAATCTTTC	TTAGAGGGAA	GCACTTCACA	12960
TGTTCCCTTC	TCAGAGGTAA	GCTTTATGAG	GCTGCAGAAC	CAGTGTCTTT	GCTCATCCCA	13020
CCAAAAGGAG	ATCTCCCACC	CATGTTCCAA	GATGGAGGTG	GGTGTGAAGT	AGGCAAAGGA	13080
TTCCTCTAAT	AAAGAGAGCT	GGCCTATTGT	AAGCATGGAA	GATCTTAGGC	CCATTGTATG	13140
ACACAGACTA	TGGATCACAG	CTCTTACACC	CTGCAGGTAG	TCAACATGGC	CCATAGCCTG	13200
GGAACCCCTC	TCTACCTTCC	CCAAAATGGG	ATCAAGCCTG	TTTCCAAGGC	CAACCATATC	13260
TCATACAGGT	TTCTGGGGTT	TACTTCTAGA	AAAGCCTGAC	TAAGACATTT	GGAGATGACA	13320
AGTACTCTCT	GGCCCGGAAG	GAGGTGCTCA	CCAACATGTA	CTCCCGGCCC	ATGCAGGTAA	13380
GGAGGGGCCA	CACCAGCCCC	TGATCCCAGT	AGTACCCATA	GCTCTGGCTG	GCAAGCACCA	13440
CGTGTACATA	GCCCACTACT	GTCTTGCTCT	GCTCTGGGAT	CTACTGGATA	GAGAGGCGCT	13500
GAGGAACACT	ATCTGGCAAG	AAAAGCTGCA	GTCACACCTG	GGACAGGCGC	ACTGAGCTCC	13560
AGAAGAAATC	TATCCTCTGT	GCTGAAAAGC	AGGCTCCATC	CCTCAGGAGC	TGTATGGCCT	13620
GTGGCTGCTA	GAGACCCCAG	GCAAGAGAAA	AGGTCTCCAT	CTCTACTGTA	GCTGCAGTCT	13680



GCAGGAGAAT	CAGTCTGCTT	CGAGCTTGGG	CCCATGTTCC	CAAGCAAGTG	ACAGCTAGGA	13740
GATAGATGGG	CTGGCTCCTA	GCAGGCTGTC	ACAGCCCTCC	AGCCTACACT	GCAGTCTCTG	13800
CAGGGCCTAA	GCATCCTTGG	GATGGGAGCC	ATCTCAGTAG	ATTGGCAGGT	CAATTGGAGC	13860
TACAGGTACT	AATGGGGTCA	GCTGTGGGCC	CCAGCACTTG	CCAGGGCAGT	GGCAGGCCAT	13920
TTTTCAAGGG	TCACTCTCAA	CAGATTCAAT	CTGTTTATGA	GAGTCAGGTA	GCCTCAGCCA	13980
GCCACAGCTG	ATTTATTTCC	TGATAACTCC	TGGCTCTACT	AGGAATGGAG	CCATCAGGGC	14040
CGTTCGGGGA	CTTGGCTGCC	TGTTCCCCAC	CCTACCACCT	ACCCTAGACA	GTGCACACAA	14100
GACCCTAGGC	TGTGCCCTGT	GGAGTGCTGC	TCCCACCAGG	ATTCTGATGG	CAAGGACTAA	14160
GTGGCAAGTG	ACAGGGACAG	GTCAGGGCAC	AGCAACAGCA	GCACAACAGT	GGGGAGTGAG	14220
GCCTGGTTCC	CAAGAGAGCT	GCTGAAACAG	GACACAAGCT	GTCCCAGTGG	TCTCTGGCCA	14280
CTACAGAGAA	GCCATGATTG	TTGCCCTGCC	CAGAGATAGC	TACTCTGACC	AAGGAGGAGC	14340
CTTGACCTCT	TTTCTCTCTC	ACGCTGCCTT	TCTGAGGAAC	TGAGCCACCA	CTGAAAACAA	14400
AGATAAACAT	GACTTACTAT	GAAGACTATG	CCCTCTGTCC	CCAGCAACTT	GCCCCAGATG	14460
TAGCTCAAGA	TCCAGCAGGG	GGCTGTGCTC	TGAGTTCTAG	GGCTATGTAC	ATGGAGTAAC	14520
CAGAAAAGGA	TGTCATTTGG	CCAGGGATTG	TGGAGCTTTC	AAAGAAGTGA	ACATCCTTCT	14580
AGGCAACAGC	TGCTGATTCC	AAGGCTGTGA	TGGCTGAAGC	CAGACCTCAT	CTAGGTTGTT	14640
CCTAGGTTGC	AGCGGCTCAG	TGGTTCCTTT	GGCTCAGGTC	TCTTAGACCT	GTGGATCACC	14700
GTGGACAGTT	GTTTCAGGAGC	AACTGATGC	AGGCTGGCAA	GCTAACAAAC	TACCCTCTTG	14760
ACTGGCATAT	GCTAGAGTAT	TGTACTGTAC	TTGTACTTGT	GGCTAGTGTG	ACCATCAACT	14820
GGGAAGAGAT	CAGAGCCAGA	GGAAATATGG	TTGGCTCAGC	CAGAAGCTGA	GGAACCTTAC	14880
GGGCTGCTCT	CCCTTGGAGG	TTGGCATCTT	GGGCTGGCCA	GGGACATGCG	GCATCCTCAG	14940
TTTCTGCTTG	TGTCTCCAGA	AGACAATTCA	CAGCCCTGGG	CCAACATGGC	CATATGTTTT	15000
CCTATCTGCA	ATCATCTTGA	CCCAGGGTGA	CTGCTCGGAT	CCTAAGGAAA	ATTATTCCAC	15060
AGCAACTCCT	CTGCATCATT	CCTGGTAGGG	ACTCAGCAAC	CATAGGCCTT	AAGGAGGAAG	15120
AGCCCTTGCA	CAGCTGCCCT	GGTGGCTAGT	CCCACAGTGC	TAGAGGCCAC	CCAGCATCCT	15180
GAGGGCTTCC	AGCCTCCCAT	GCCCAACAGA	GGCATAGCTT	CCTGAGCTGT	TGCGAGCATT	15240
GCCCTCATGA	ATGGAGCCCG	GCAGCCCTAG	GCATGACTAG	CATGCATCCT	GAGCAGGGAA	15300
GGGCTCTGGT	CATTACATGC	TGTCCATGGC	AGCTGCTGAG	AACCCCTTAA	GTAGGATGAC	15360
CCTGGCCCCA	AGAATCTGGG	GCTTTGATCA	GCTGCCTGAA	GCTGATAGGG	GAGGTGTGTA	15420



TCAACCTTGC	CATGGGCCAG	GCTTGGGTCT	CAGCACCTAG	CCGACCCAGC	CAGGCTTAGT	15480
CCCACTCTCC	CTCCAGATGG	CACTGTACTT	CTGCTCTGGG	ATGCTGCAGG	ACCAGGAGCA	15540
GTTCCGGCAT	TATGCTCTCA	ACGTTCCCCT	CTACACACAC	TTCACCTCTC	CCATCCGCCG	15600
CTTTGCTGAC	GTCATAGTGC	ACCGCCTCCT	GGCTGCTGCT	CTGGGTAAGG	GACATGACTC	15660
TGGCCTGGGA	AGACCTTTGC	TGGTCGAGAG	TTACCCACTC	TCAGAGTAAG	TGACCACATT	15720
ACTGTTATCA	TGGACATGCC	GAGGGACAGA	GAAGCCTAAG	TCTGAACACT	GTCGATCCAC	15780
ACCCAGATGA	TGGAAGCTTT	AGTGAGACTT	ATTGCAAGCG	CGGGACCATA	TATGGTCCCA	15840
GAGCCTTGCC	TCAGCACACA	ACCGTCCTTA	TCCCCATACT	AGCAACCCTG	GTCGCCCTCT	15900
CCTCCAGGCT	ACAGTGAACA	GCCAGATGTG	GAGCCTGATA	CCCTACAGAA	GCAAGCTGAC	15960
CACTGCAATG	ACCGTCGCAT	GGCTTCCAAA	CGTGTGCAGG	AGCTCAGCAT	CGGCCTCTTC	16020
TTCGCAGTTC	TAGTAAAGGT	GAGTGTCCAG	CCTGGCCCCCT	TCTTCTTCCC	CTTTCCTGT	16080
CCTCCGATGA	ATGGAGCACC	AGTGCAGGTC	CTCCCTGGGA	GGATGCCACG	ATGCATTGTT	16140
CCTACAGGAG	AGTGGCCCCC	TGGAGTCCGA	AGCCATGGTG	ATGGGTGTCC	TGAACCAAGC	16200
TTTCGACGTG	CTGGTGCTGC	GCTTTGGGGT	GCAGAAGCGC	ATCTACTGCA	ATGTGAGTAT	16260
CCCTGGTATG	AATGGGAGGC	CTGCACCTAC	AGGCAAAACC	AAACCCATTT	TCCCGCCTGT	16320
GTCTAGTTCC	TTGTTGGGGA	AATATTCCCC	TGGTCCAGAA	TATCCCATGA	TAGTTTCACA	16380
GGTGTAATG	GTGGGATTCA	ACTGAGCTCC	CTTCTGTCCC	TGGCCATTAG	CTATGCAGGG	16440
CCCACAGACT	GCATCCTATA	GCAGTGAGTT	TCACTGGCAT	GTGGCAAGAA	AGGGTCCAGA	16500
CCCCTGAACC	CAAGTAGGCC	TGCCCAGGAC	AGGGCCTCAG	GCCAAGGGTC	AAGTCTGAAC	16560
TCTTCCTTAA	AAGCCCAGGC	ACTCAGAACA	TAACCAGGAT	GGCAGGGTGT	GGGACCTGTG	16620
ATGTTCTTAT	AGAAACATGC	AGAAGGGGAG	GCCAGAGGGT	AGCCAGCACT	GCTCTGGACA	16680
CTGTGTCCCC	AAACAGAAAC	AAGAGGCCCA	TCCTGCCTTG	GCTTCTTCCC	TGGATGACAG	16740
TTTATTCAAA	GTCCTCTTGG	TGCCTTCTGT	AATGTCACTT	GGGGGGCTTT	GCTTTAGCTG	16800
CTCTGTGGTC	ACCAAGTCAC	CACCTGGCTC	CTACCCCTGG	CTTTGAACTT	CTTACATACA	16860
CTTGGGGAAG	TGTGGAACCC	TGCACTGGAA	GAGACACAGG	ATTCATGAAA	GAGGCAGAAC	16920
AGGAAAGGGC	CAAGTGCAGC	TGGAAC TACC	AGACACCTGT	AGTTACCTGG	CTCTCAGCCT	16980
GGTGGTCAGG	TCTATCACCA	ACAGCCTAGG	CAGATCTCTT	CTCTTTGCTA	CAGTCACCAC	17040
CCTCCCACAT	TGTCCCTTGG	AATTGGGTCA	CCTTCAGGTT	CTACTTTGAC	CAAAGGTGAC	17100
TTAGCAGAAC	CTCCTAAATC	TGGCTGAGGT	GGACCAAGGA	TAGGGGGCTG	GGGGATGTCT	17160



CTGTCCAAGC	AGGCAGCTAC	AGTAAGGCAG	CCGGTACAAA	GCTCCCTCCA	GCCAGTCAGA	17220
AATAGGCAGG	CAGGGCAGAA	GAGGTGTCTG	AAGCCCATAG	CCTGAGGCTC	CGGTGTGTCC	17280
CCCTGCCCCC	AGGCACTGGC	CCTGCGATCC	TACAGCTTCC	AGAAGGTGGG	GAAGAAGCCA	17340
GAGCTCACTC	TTGTTTGGGA	GCCTGATGAC	CTTGAAGAGG	AGCCAACACA	GCAGGTCAGT	17400
CCCCTGCTGT	GTCCCTAAGC	CTACCTCTGT	CTCAAACGTG	TGCCCCTAGG	TCCTCATCTG	17460
CCCTCATTTC	TCCCCAGCAC	CATAGGTTCC	CCTGTGGGAT	TCCACCAAGC	CCTGGCTTAG	17520
ACTGCCAGGT	TCTATATGGG	AACACCCACT	ATGGCAGTGG	TTCTCAACCT	TCCTGATGCA	17580
GCGACCCTTA	ACACAGTTCC	TCATGCTGTG	GTGACACCCT	TCCCCAGCC	ATTAAATTAT	17640
TTTCGTTGCT	ACTTCATAAC	TATAAGTTTG	CTGCTGTTAT	AAATCAAATG	TAAATATTTT	17700
TGGAGATAGA	GGCAAAGGGT	CTCGAACGAC	AGGTTGGGGA	CTGCTGCTCT	ATAGGTAGAT	17760
AGGTGCTATT	CCTCTCCCCT	GAACAGAACT	TTTCAGAAAT	TTTGAGAAGC	TGATAAAAGC	17820
TTCTTTTATC	CCTCTTGTTT	CAAAGGCTGG	CCCAGCCCAG	CTCGGCCCGG	CCCAGCCTGT	17880
TTTCTTGCTC	CTCGTGAATG	GTCAGTGAAT	AACAAATGTC	TACATAGTGC	CATTTAGCCT	17940
ACTGGTTTTT	CCCAGACCCA	ATGAATCCCA	TTTACAGATA	GGCGATAGAG	GCTCGGGAAG	18000
TTAAGTGAGC	CTCAGTGGTC	AGTTGGCTTT	GATTGCAGGC	CCTCACCTGC	CCTGTCCTCT	18060
CCTGTTCTTG	GCTCTGCTAC	AGGTCATCAC	CATCTTCAGC	CTGGTGGATG	TGGTCCTGCA	18120
GGCAGAGGCC	ACAGCCCTCA	AGTACAGTGC	TATCCTGAAG	CGACCAGGCC	TGGAGAAGGC	18180
GTCTGATGAG	GAGCCTGAGG	ACTGAATGCT	AGCCCAAGCC	AGGCCTGTGC	CTGCCCTACC	18240
CTGCTGGCTT	TTAGGAATAG	GACCTTTTGA	CACCAAAGGG	GATTTTAAAT	TTGGTTTTTA	18300
ACAACCTCAGG	GGTTTGTTTT	TATTTTTATT	TTTCCTTTTA	TTTTACTTTT	GCAGCTCAGT	18360
TTTTAAATGA	ACTGGAAGGT	TAGGGGTCAG	GGCAGGGGAT	GCTGAGGCCT	GGCCTGTGCT	18420
TCCCTGAGCA	GAGAGGATCC	CAGTCCTCCT	GGGCAGGCAG	CCCCGCTTCT	ACCAGGCGAC	18480
CCACTGCCCT	TCCCTGCCCA	GGAAATGGGG	GGTTTCAGCA	AATCAGTGTC	ATGGAATAAA	18540
ATCAAGTGTG	AATTGCTGTC	TGTGTAGATG	CCATGGGCAA	GCATGGCAGC	TGGGTGGCCT	18600
GTCACCGAGG	GCAAGGGGCT	CCCTAGAATC	CACCTCACAG	CTGAGCTGGG	GTCATCAGCT	18660
CAGGACCTTC	CTGCCAGCTC	CAGGGTGATT	CACGAGCCAT	GTGTGGCAGA	TTGATGCTGC	18720
AGCCTCCTTC	TAGCTGATTA	AAAATGTAAT	TAGTATGCAC	AGTAGGGAGC	TGCCAGTCAC	18780
CCTGTGCATG	TGGCTGTGGC	CCTCCCTCCC	CGCCCTTCCT	CTCTGTTGCC	AGCCCATGGG	18840
ATGTGGGGAG	GTGGGACTAC	CACCTCTCTT	CTTATATATC	ATAGGCCAAA	GCTCCCAGGA	18900



GCCCTGTTCA	CAGCTATGCT	ATGAGTAGGT	ACCTCAATAC	CTGCAGTTTC	AAACATGTAC	18960
CCTAAAAGGT	AAAGGCAGAC	CTTCCAGAGG	GCAGGAGGAC	TTCAAAACAG	ATCCTACCTG	19020
ACCCAGCCAC	CTGCTTAGCA	TCCCAAGTAC	TAGCAATTCC	TACCCTTCTG	AGCACTGGGC	19080
AGCCTCTTCC	CTAGGGAAC	GGGCACAGTG	TATCCTCCTT	TCACCAGACT	GGAATAGTAT	19140
GAATTGGCTT	CAAAAGCAAC	TAGAATCTAG	GATGAAAACC	AAAGCAACCA	AGGCCCTGTT	19200
CCCCAGTGCT	GTTCCCTGTG	GCATCAGGAT	TAACAGACCC	ATCTGATATG	GTTATGGTGA	19260
TTTTCTTCAA	AAAAGATTCT	GTGGAGTCCC	CTGGCAGGTT	CCTTGCACTG	AGTGAAGTGGC	19320
ACAGCTGCAA	GGATATCACA	GCCCTAGGAT	GGGCTGTTGT	CTGAGGAGAG	CCACAGACAC	19380
GCCCCACCTG	CCCTGGGCTC	CTTGTCAGCC	TCACACAGCC	TTCAGCTGCC	TGTCCTCCCA	19440
CCCCTTAGGT	CTCCCTTCTG	CTCCCATTC	CAGACCAGCA	TATCTGGATA	GGCAGAGCAG	19500
TGATGGATGG	TGGTTTAGTA	TCTGGGTAAA	GAAGACTCTG	GTGCTTTGCC	AATCCTGGAT	19560
CTCTAGACTA	AAGGCTCATC	CCACAAATCT	GAGGAGGAGC	TAGCTTCTCT	GCTGGGCCAA	19620
ACCCGGGCTT	CCAAGACCTC	CTTTCAGTGC	CTCCTTCAGA	ATCCTTAAGG	AAGCTGTGGC	19680
TCGAGTACTG	GGTTCTCTCA	AGACACAGAG	GTGGCTGAGA	CACGGCCTCC	CCAACCCTCG	19740
TGAGGAACAG	CTTACCAGTC	AGTAAGGAAA	GTTTTTGCAG	AGTGAACGTG	CTTAGGAGGC	19800
AGGCACTGGA	CTAGAACTT	CTATAACAGG	CTTGCTCCAC	CCTCAGGTTG	GACATCATGT	19860
TACTGAGAAC	TCTGAGCCAT	AGCAGTCCTG	GGTTGCCCTA	ACCTGTCTGA	CAAATGGAAG	19920
TCTCAGGTCT	CCATCTGAGG	TGGTGCAGCC	AGGCCGCCCT	GGCCAGGACT	TGAGCCACCT	19980
GTCCTCTGTT	GCCTCCCAGT	GGCTCTGTCA	TCTTCCCACA	GCACCAGCTG	AGTCACTTCT	20040
CTTTGTGTTT	GTTCACCCAG	CACTGAGTCA	GAGAAGTATG	AGAACGTGTG	TCCACACACC	20100
ACTCAGTGTG	GCAGTTGGCA	CCGAACACTA	AGGGCACTGC	TGGCAGAAGA	GATGACAAGA	20160
AATAAACGAA	GTAAGTCACTC	ATCAGCTATC	CAAGACACCT	GCCTGCACTA	TAGGCTAAAG	20220
CACAGGGCAC	AGAGCAGCTC	ACTGGCTTTT	CCTCAGTGGC	CTGTCAGGTT	CACATGGAAG	20280
GAAGACAGAC	ACAATCTCAC	TCTGATTGGG	GTCTCAAAAA	GCTCAGAAGC	AGGCAGTATG	20340
TTCCCAGGGG	AAAATGGAGC	AGGTTGTGGG	TCCAGCATGG	ATGAGAAAGT	TAAGTATTAA	20400
TTAATGGTTG	TAACCTGCCC	TCCTGGGGAG	AGAGGCTGAC	ACCCTGCACA	GTCCTACTTA	20460
GCAAAGAGCC	TTGGAAAGGA	CTTCAGTGGG	CCCAGGATGG	CAGTCCACCG	GAAGCTGGAG	20520
CACAGCACAC	TGGAGGTATG	GTAAGAGGGA	GCTGGTGCCA	GGCAGAGGCA	TCCCAGATGC	20580
ATACCGCAAC	AGCCAGTGAG	GATACCCACT	GCACCACCAT	GCCAGCTAGC	CACTAAAGCA	20640



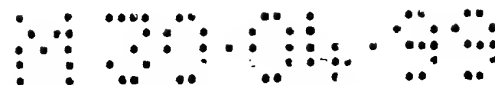
GCCAGTGAGG	GCAGTCCAGG	TGAGAGGAGG	AAGGCCTGAG	AGGAGAAAAA	AAATATCCAA	20700
AATCCTGGGG	TGGGTGGTGT	CCCAAACTG	AGGCAGCATA	GGCACAGTGG	GAGCAGCAGA	20760
GACCTGCAGT	GGCTCCTGCT	GGGAATGGGG	CAGGCCTGTG	AAGGAGAGAG	GGCTGAGCCA	20820
TAGGGCACTG	GTGACTCAGT	GAGATGGAAA	GAGGGACCAA	GTGTAGAACA	GCTGGACCAT	20880
GAGAAGAGAG	CATGCAGGGC	AGTTCAAGAA	CCTTAGAAGA	GGCCATGTGG	GCAGAGTGGG	20940
GCTCCAGAAG	AGGGTATTGC	AGTCAATGGG	AGCTAGGAGC	CTGGAGCCAG	ATCTCCCTCT	21000
GTGAAGGTTA	TTGATTATCA	GTTTCTGAAG	GATACAAAAC	ATCCACTCTC	ACTACCTCCC	21060
CAAGACCAGC	AAAGGCACCA	ATGAGCTTGT	GTTCAGGGAT	CCATTGTGAG	GGGAAATGGG	21120
AAAATAAAGG	AGGACGTTAC	CCTGGTAGCT	GAGAGTGAGC	CAGCAGTCCC	TGTTAGACTG	21180
GAGAAAGGCA	GGTACGAGGC	CATCCACAAA	GAATGCTGAA	GCACCGAGCT	GCAGTACTGC	21240
ACAGCATCCA	ACAAGGCTGG	GCTGCTCTGG	GCTGGGGGTG	GAGAAGGATG	GCTACAGAAG	21300
TCAGTGTTGC	CACTGTAGTA	AATAAACTGA	CCTCTTCCCA	CACCAGCAGG	CAAGAGAGCG	21360
ATCATCGGAG	AGTCACCAGG	CCTGGTAGAA	TCTCCTGTGA	TAGGACCCCA	TGAGATGCAG	21420
CAGAGGGCTG	CTGCAGGATC	CAGTCAGCCC	TCAGGCCTTC	AGCAGCCAGG	CAGGAGATTG	21480
AAAACATCTT	CTCCGGGGCC	CTCCTGTCCC	CACATGAAAT	ACAAACTTGG	CAGCAGAGTT	21540
TCCCCAGTGA	GATCCCAGCC	AGGCTTCTCA	TGGGGAATCA	GCCTGCCAAG	TCCCTAGGGT	21600
ACTTGGGCTT	CTAGTCACTT	TGTGAGTCCT	ATCTGTAAAT	AAAGATAACC	AGGGAAACTT	21660
CCTTTTAAAA	GGAAAATAGG	TCCTATGGAG	AAAACAGATC	ACACAGAGAA	AATGAAGTTA	21720
TCACTGACAT	TTTCAAGGAA	ATGAGAGCCA	TGGAAAAACA	AGGACTAGAT	GGCTAGACAC	21780
CAAAGAAAGG	GCTGGTGATG	TAGCCCAGCC	AGTAAAGGTA	CCAGGTGCTA	AACCTGCCAA	21840
CACGGGTTCA	GTCCCAGGGC	TCATAGCAAG	AGCAGCCAAC	TGTGGTTGCT	ATGTAATGTC	21900
CATAAGGCGT	CTTTGGAGTG	TTCAAAGTAT	CTAAGCTCCC	ATGAAGGCCA	TCCAGCTGGC	21960
TGCTTGGCTA	ATATCCTTAA	ACATCCAAGG	TTCCAGAGAA	GGATATAGTT	ACAGTTAAAT	22020
CCCCCTGGCT	CACAACATCT	TAAC TTATTT	GAAAAA AAAA	ATATCTGAGC	ATGGCAGCTC	22080
ACACCTGAAA	TCTCAGCATT	TGGGAGCCTG	AGGCAGGAGG	GTTGCCATGC	ATTGGAGGCC	22140
AATCTGGGTT	ACACAGTAAA	TACTAATCAG	ACTACGTACA	AGACTATGTA	GATATACTAT	22200
GTAGCAAGAC	TGTCAGAAAG	GAAAAATAAA	CATTAAAGAG	GTAATTAGAG	TAAACGCCCA	22260
CCATTAACTG	TAATGGTATT	TAATAGTGTT	CAACCCTCAA	CCAAATGTCC	CTGGGAGGAG	22320
TTGGATTATT	TTATGTCTCA	TACACCTAAA	CAGTAGCATC	AGTGCGCTCA	GGATTGAGGA	22380



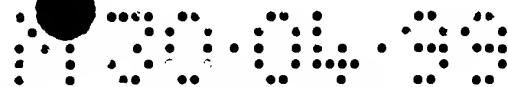
GCAGGCCAGC	ACCACCAGGG	GTGAGAGGCA	TCCGATCTAG	AAGATCCCTG	CCTGAGGTAG	22440
CCGGTAAGTG	AAGTGGCTCA	GAGAAAGTCA	AGTCACGGAC	AGACTCCAAG	ATTAGACTGA	22500
CACTAAGTGC	ACTGAAAACA	ACCCTATCTG	ACAGTAAGGA	ACGTATTGGG	TATGAGTGGG	22560
GAAGCAAGTA	CAAGAAAGAA	AAGCCTTTCC	CTGGTCTTTC	ACCTGGCACA	TCTGGCAACA	22620
GCAGTACATC	CTAAGATAAA	CACTGAGTGA	GAATCTACAA	ACTGCTCTGG	GGCCATATTG	22680
AGAGGATGAG	GAGATGGGAC	ACATGAGTAG	CCAGTTCACT	CTTCAGTGGA	AGGTTCTGGG	22740
GAGCTAAAGG	TGGCTGCAGA	TTCATTGCCT	ACCCACCACC	ACCACACACC	CTGTTCTTGT	22800
CCTTCCTCTT	GAATCAGAGC	AGAGTCTTCA	GCTGCTGAGC	TCAGATACAG	CGGAAGTGAT	22860
GTTGCACTGT	CTCCGGCCAT	GCTGAGAGTG	CCACAGCAGA	GCTGTGAGAA	AGTTTGGGCT	22920
CCCTCGTACT	CCAGCTCAGA	GGCATCTTAG	AGATGCATGC	CCAACCCCCA	CAGAACCACC	22980
CAGTGGTGGC	CTTGTGGAGG	AAACACAAAG	TCTCCAGAAG	ACCCCTTCCA	AATTACACAT	23040
TTCTATCAGC	TTTAAAAAAA	AATGTTGGTT	GTTCAGGGAT	AGTTCATGAC	ATAATATTAG	23100
CAGAAAATGT	CAGTAAATAC	AGCTGAAAAC	TGGAAATGAA	GGGCTGGAGA	GATGGCTCAG	23160
CAGTTAAGAG	CACTGACTGC	ACTTCTGAAG	GTCCTGAGTT	CAAATCTCAG	CAACCACATG	23220
GTGGCTTCAC	AACCATCTGT	AATGAGATCT	GATGCCCTCT	TCTGGTGTGT	CTGAAGACAG	23280
CTAGTGTTCT	TACATATAAT	AATAAATAAA	TCTTTGGGCC	AGAGTGAGTG	GGGCCAGAGC	23340
AAGTGGGGCT	GGAGTGAGCA	GAGGTCCTGA	GTTCAATTCC	CATCAACCAC	ATGATGGCCC	23400
ACACCATCTG	TTCAGCTACA	GTCTACTCAT	ATACATAAAA	TAAATCTTAA	TAAAAAACTG	23460
AAAAAGAAGA	AATGGTTGTT	TTCATTTGTC	TGTTATTCTG	AGAGGTGTGG	TTTTTACAAA	23520
TAGTGGTAAC	TATAAAAAAT	TTAAAACCCA	TGCAGATTGG	GGGTGGACTA	GGGAAATGGC	23580
TCAGTAAATC	AAGTGCTTTC	CACACACAGG	AGATGCACTG	GAGCTCTGAT	CCTCTGAACT	23640
CCTACACAAG	CAGGCGGCC	TGGCAGCTGC	CTGACATCCC	CGCACTCAGA	GGCCCTGGTG	23700
AACTGACTAG	CTAGACTAGC	GGGACCCGTG	AGCTCTGGGC	TCAGACAGAG	ATCCTGACTA	23760
TAGAAAGTAG	AAATCAACCA	GGGAAGGGGT	CTGCCTTCAA	CTTTGGGATG	CCACATTCAA	23820
CCACATGCTC	ATGCACACAC	ACGCACGCAC	GCGCGCGCGC	GCACGCGCAC	ACACACACAC	23880
ACACACACAC	ACACTAAATA	CCAAGAGGGG	ACGTGGTTGC	CTCCAAGATG	GAAAATGCAT	23940
CTAGGAGCAT	GAAGTGCTCT	CCCATTTTGT	TTTAATAAAC	CTGCCAGATC	CATTTGACAC	24000
TTTACATCTG	TGTATAATTT	CAATTTAAAA	AACTAAAAGT	AGGGGGGAAG	GCTGTTTATA	24060
TTTAGCCAGA	ATGGATCCAC	AATTGGTCTA	AAAGCTTTCC	TGTACATTCA	GCAAGGAGTG	24120



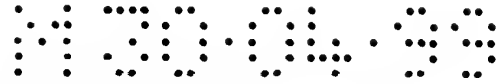
TATTAAACAA	TCCATTATTC	TAGTAACTAA	GATAAAATCC	CTGCTGACAG	GCACCCTGGT	24180
ATTCCCAGAC	CATTA AAAATG	CTTCCATAAA	GTCTGCTTAA	AGACACAGGT	AGCAGGCCAG	24240
GTGGTGACAC	ATCCTGGCTG	CCTCAGCAGA	CCTTGCAGGT	CTAGGTGTGG	AGCCCAGAGT	24300
GTGGGGCAGC	CCTGGGGCAA	CACAGGCAGA	CCTCTGGAGG	CCTGCGGAGG	TGGCATGGCA	24360
GACGACACTG	TAGGCAGCTT	GCAGAAGAGC	TGGCCAGGGG	CCTTAAAGGA	CATCAGCTAA	24420
AGGCCTCTGT	GGACCGAAAG	CACAGGCTTG	AGGGATTATT	TGGAGTCGGG	GTTGGGATGA	24480
AAGGAATTGA	CACAGATTAA	AGAATCAACT	CCACTCTGGT	GGGTGCCAGA	ACAAAGGTGA	24540
TGCTTTGTAT	AACGATGAAG	AAAGTTCTAG	AACTAGGGGG	CAGCTCCATG	ATAGAACACC	24600
TGCTTAGCAG	GTAAAAAGAG	TCAGGTTCAG	TCTTTGGCAC	AACCCCTTA	AGAAGGAAGG	24660
TTCTAGAGAA	AGGGGTGTTC	TGGACCTGAG	AAAATTAGCT	TGAATTTGCA	TATAAGTAAA	24720
TTATGTTTAT	AAGTTGAAAC	TCTTACCGTG	GCCCTGGAGA	GTGGCTCACT	CAGTTAGTTA	24780
GCTGCTCTTC	CAGAAGACTC	AGGTTTGAGT	CCAGTGACTC	ACAGCTATCC	ATAACTCCAG	24840
TCCCACAGAG	ATCTGATAAC	CTCTGGCCTC	CTCAGGCACG	CACCAGGCAC	ACATGTGATA	24900
CACAGACATA	CATACAGGCA	TACCATGAAA	ATAAATTTTA	AAGAATTAAC	TGTAACCAGG	24960
TCTGTTAGCA	CATCCCTGTA	ATCCCAGCTG	CTCAAAGGGC	TGAGGCAGTA	GGAGAGCAAG	25020
TTCAAGTCTG	GCTTTGGCTA	CAGAGCCTGT	GAGTTAAAGC	CCAGGCAACT	TAGCAAGACC	25080
CAGTCTCAAA	ACAGAAATTA	TAGGCAGGAG	GTACCTGGAG	CCATAGCTGA	GGATGGGTAC	25140
TGGCCAGGCC	TGTGTGAGTT	CCCCAAGTTC	TATTCTCATT	CCTGAAAAAA	AAAAACAAC	25200
AAAAAAAAAA	ACATAAGTGG	TCAGTTAAAC	CTTAGGATAA	GATAATCTCT	TTGAACCTGC	25260
TCTGCCTTTT	TGTGAGCTTT	TATGATTATC	AAGGGTTTCT	TTCTCTAGTA	TATAAAGCCA	25320
TCTTAGGGGG	TAAGATCTAT	TTAAGTCATT	TATTTTACTT	AAAACGGTCA	TTTTACTCAA	25380
GCAGGTTCAT	GAAC TTCACT	GTGTTCCACA	GTGTTCTTAA	ATTGTACAGT	TCTGGAAAGC	25440
AGTTAGCCAA	ATACCAAGAA	AATGAATGCA	GAATAGAGTG	AGGAACAAAG	GCGGCCCTTC	25500
AGCATATTTT	ACCTTAATAG	ATTTTCCAGC	TAATAAGACT	GCTGCTGGAG	GGAGAGTGTC	25560
CTCCCGGTGC	TCCTGACACC	AAGTCACAGA	AGAAATTACC	GAATGCGGCA	CTGGACACCT	25620
AGGACTTTGC	ATTCCTCCAT	GCCCAGAGAA	GCAGGTATCA	CTCAGAAGGA	TGACAGGGGC	25680
TGGGGAGGTG	ACTCAGCAGA	TAAGGCACTT	CCACAAAAGC	CTGATGACCT	GAGTTCAATC	25740
CCCATCACCC	ACTTTTTTTT	TTTAAAGAGA	GGAAGGAGAG	AACTGACTGC	AGTTGCCCTC	25800
TGACTTCCAT	GTGCTCCCCA	AGGCGAGCAA	CACACCACAT	CATACACATC	ACAATAATAC	25860



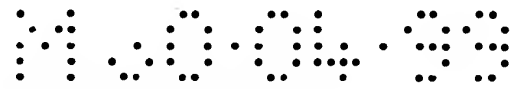
ATTTTAAAG	GATGACTTTG	AGCTACACCT	GCCAACTGTC	CCTGATGCTG	CCACCACTAC	25920
AACTAGACAG	AGGAGGTCTT	GCCTGGTGGG	TAAGTGAACA	GTCAAGGGTG	CCCACGGAGA	25980
GCCACTTCTG	CCAGGCCAC	TCCTGAACTC	CTAGGTCCTC	ACGGGCTCAG	ACCTCTTGCC	26040
CTCCGCTGAA	GCTGCAGAAG	GGACTCAGCT	GTGCACTGTC	TCCTCCCCCA	GGGACCATGG	26100
GGCGTGGTGA	GGGAAAGGGG	ACTGTCTCTT	GCCTTGGTGG	TAGATCAGTC	TCCTTCCTGT	26160
TCTCACACCA	GAGCCCAGGG	ATTGACTCAG	GTGATGAGAG	AGTGGAGAAA	GGATCTACAC	26220
CCAGCCCCCC	TCTAAGACCC	CATAGCAGCC	CCAGGACATA	AGTACAGAAG	AGCTGGGCTG	26280
GGCTATGCAT	TTGCTTTATA	CATTTGAGTC	AGGAAGGTGG	GCTTATGGTA	CACAGCTGAG	26340
CAAGGAGGCA	GATTTAGCTC	ATCTTTATAA	GAGGTCTCTG	TAGGGGAGCA	GTCTTAGGCT	26400
GCAGTTATCC	CAGAGGAGGA	AGCTGATAGC	TTCTACATGG	ACTGTTAAAA	TTTGCATTCA	26460
GACCAGGGAA	AGGCTTTGCC	ACCCCTCTGA	GCTTCACTGG	GGAAGGCTTC	GCCACTCCAT	26520
GGGCCTGATG	CGTTGGAATC	CATGACAGCT	CAGCCCATGT	CAACAACACA	CATTCACTTA	26580
GGGTTTCATC	TGCTCCTTTC	ATGTAACACA	AGGCTGCTTC	TGCTACGTGT	GGGGATTGTTG	26640
AGAGTATATT	TCTTGCTGGA	AATGAATGAT	CAAAGCAAGG	CCCCACCTCC	TAGGCTCTAT	26700
CAGGATAGAA	GGGTCACTAC	CAGAATGAGC	CACCTCCTCA	CTGACGGTTG	GCTCCACTTG	26760
CAGGCCTTCC	AGGATTCCAA	GACTTGGTTC	TTTGTTCTGA	AGCTCAGGGT	ATAGCTTCCT	26820
CTACCTCCAC	ACACAGCCCC	TAACCTTCA	GTGCATAGTG	AACCACTAAG	ATCTCCCACT	26880
ATGTCCCCAT	AGCAGCCCTG	GAGTACAGGT	CCTGTCTCTT	GCCCATTCTC	AGGTGAGAGA	26940
ACCTAGGCTC	AGAGAGATGA	CACTTCAGAA	GATAATCAGA	AAATGGTGGA	GGTGATTGGG	27000
AGCTCAGATC	CAAAATGCAC	TGCATTTCTT	TATTAGATAT	TTTAAATTCT	AACGGTGTAC	27060
CTGGGTGTTT	GGGCTGCATG	TGTGTCTGTG	CATATCACCG	CTGTGCCTGC	TGCCCACAGA	27120
AGCCAGAAGA	GGGTGTTGGA	TTTCTTTCTT	TCAATTAGTA	CTTCTCAAAA	TTCAACTATT	27180
CATGCATCAC	TTTAATGATT	TTTTTTTTTT	TGCCATAGCC	ACATAATGGC	CTGTGGTCAT	27240
ATTTATTTAA	TGTTTTTCAT	TAAACAAGCT	TAGGCCTTTC	CTTGAAATAA	TTAGAAAGGA	27300
AAACTTACAG	TTACCAAAAA	ATAGAGGGCC	AGCTGGGGGT	TTAGCAAGAG	TTGGTACAGT	27360
GTTACCTCG	TATGCACAAA	GCCCTGGCTT	CCACCCCCAG	TACCCAGAGC	TTGGGAGAGG	27420
AAAGGCAGGA	TCAAGAGTTC	AAGGACATGG	CCAGGCATGG	TGGGGCATGC	CTTTAATCCC	27480
AGAGGCAGAC	AGATCTATGT	GAGTTTGCAT	TCATCCTGGT	CTGCAAAGTG	AGTCTTGGAC	27540
AGCCAGGGCT	CTGTTACATA	GAGAAACCCT	GTATCGAAAA	ATAAAAAAAC	AAACAAACAA	27600



CAACAGCAAA AGAGCTTAAG GTCATCTCTG GCTGTATAGC AAGTTTGAGC CCGGCTGGGC 27660
TATACAAGAC CATCTTAAGA GGGAGGAGGA AGGGGAAGAA AAAGAGGAAA CAAGAAAGGA 27720
GATAAAAGAA GGTGGGGGGA GTAACCAGAA CGCATTATAT AAATGCATGA AATTGTCAAA 27780
GAACTAAGTT AATTAAAAAG CAGGAAGACC ACCATCACCA GCCTCGAGTA GAAGGCAGCT 27840
GTGTATTCTA AGCCTGCAAA TAGCAGTGTG AGTCTTTGCT CCGGGGCTCT GCTTCAAAAG 27900
AGATGGTAAA GTTAGTACAA TGTTAGAGAA TTTCAGGAAC CAACTGCGAT CCTTTCCTCG 27960
ATATCATCAA AGGGGTGGAG AGAGAGACCA ACAACGCTCC ATAGCACAGG CCCATCACTC 28020
ATGTGCCTGA GAAGCTGGAG CCAAGGATCT GTCTCTTCAA GACTCCATCT CAATAATGGT 28080
TCAGTGACAT TTTATGCCCA TTGGTGATAG CTAAACTAGC CCCATTTTAC CTAAAAGCCC 28140
ACACCTGGCA CCGTAGTTTG TCCTGTCTTG CAAAAAATGC CGGTCAAGAT GGAGATAAGA 28200
ACCGTGGCAG GAACAGATGC ATCTGATCTC AGTCACACTG CCAACCTATT CCTTCCTCCT 28260
GAGGCAGCTC ATGCTGAGGA GTGCTGGCTA GCACCAGTGG TACACAGCTG AAGACCATGA 28320
CTCGCCTTCT CCCAGAATTC CCAGCAAGAG GCATTGAGCC CAATAAGTCC CCCCTCCAGC 28380
CATGACTAAT TTTTGACAGT GTCCATCTTC TGATAGCCCT TGAAGGTAAC TACAGCTTCT 28440
GTGAGTTTAT GATTGTGATG ACTGTGGCAT TGTCAAAGGA TGGCATTTCG AAGTCCTCTC 28500
TGCCTTCTGG CTTGCATTTT CTCTTCTTCC TCCCCACCT TGTTCCCCAA GCCTTAGGAG 28560
AGTGGCATCT GTGTCTTGTT CAGAGCTGAG CACTCAGCCA CCATTTCTTC TCAGTGCCTG 28620
GGCCTCACAT GCAGTCCTTG GGCAGTGGTT GGTGGGTCCA GTAACAAATA GGCATGTCTT 28680
GCCTAGCAGG TCTTATCTAG CTCTGGTGGG TTTCCAAGCA TGTAGCAAGA AGAGTCTGCA 28740
CTGTTTTGGG AGTCTCTGGA GCATCCCTGA CCAATGACTG ACATGGAAGT GCTCCAAACC 28800
TCCTGCTTCT GGGGTTTCTG TTTAGTAACC CACAGCCTCT AGGAACAGTG TTATCCAGAC 28860
ATGTAGGGTA TCTCTCTTCT AATGTGTGCG TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG 28920
TATAATTGTG CTACAATATA GTAAGTTTAC ACACTTGTTT TGGTTAACCA CCCCCACCC 28980
ATCCCGTCTC CCCACTTCTT TCTCTAATTA AATCTTTCCA CTCCAAAGAG CATTACTGCT 29040
ATTGCAGAGA ACATGGGTTT GCTTCCCAGA ACCCACTTGG CAGCTTACAG CCATAGTAAC 29100
TACAGTTCTG GGGAGTCCAG TACCCCTTC TGGCCCCTGC CTGCACCAGA TACACACACA 29160
CACACACACA CACACACACA CACACACACA CATATCATAC ACTTAGATAC CTGCAGGCAA 29220
GACATTTGTA CATATAAACT AAAAATAAA TCTTAAACAA AAAAAAATT TCCACTCAA 29280
GTCTTCACCC TCTCTGTTTT CACTTTATCT GTGTCTTGCT ATCCCTTCTC CCTTAAAGGG 29340



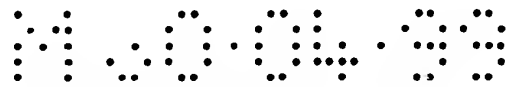
AAGAAGGACA	GAGGGAGGAG	GGAGGGAGGA	GGAAGGGAGA	GAGGGAGAGA	GAGAAAGAGA	29400
GAGACAGACT	CCTAGTTTCC	TGGCTTCCAC	AAGTGCTCCA	AGGTAAGCAT	GCATAACTAA	29460
AGAATCAAAG	CTAAGTAAGG	GCTGGAGAGA	TGGTTCAGTG	GTTAAGAGCA	ATGACTGCTC	29520
TTCCAAAGGT	CCTGAGTTCA	GTTCCACAT	GGTGGCTCAC	AACCATCTGT	ACTGAGATCT	29580
GGTGCCCTCT	TCTGGCCTCC	AGGTATACAT	GCAGGAGAAA	TGCTGTATAC	ATGATAAATA	29640
AATATTTACA	AAAAAGAAT	CAAAGCTAAG	AGCCATATGT	AAGGATGTAA	CAGCATCTTT	29700
CTGGGCCTGA	GCAACACTAT	ATATATTTTT	CCAGTTCCAT	ATGTTTACCT	ATGAATAAAA	29760
TTCATAAGTA	TATATGCTTT	GTAAAAATA	ACAAAACATT	TCAGGATAGC	CAGGGCTACC	29820
CAGAGAACT	GTCTTTAAAT	AAATAAAACA	AAACAAAACA	AAACAAAACA	GATACCAAAT	29880
CCACAAGCAG	TCCAATCAAT	ACTGAAACGC	TGGTTTTGCA	AGCTACCGGG	GTTTTAATCA	29940
TCTTAACGTT	TCTTTCTCTT	TCCATCTTTC	CACTTCTTTC	CTGCCCTTCT	TCAGCTTGAG	30000
CTTTCCTCGC	CACTGACGTC	AGCCTTGTC	TCCTCACATC	TCTCTTCCCA	CTGCAGGCCT	30060
CATCCTCGAA	CCTTCCTCTC	ACCCTTCTCA	GGCTCCTCTC	CCCTCACCAT	ATCACCCACA	30120
GCATACCCT	TCTGCAGCCC	AGTCAGGACC	TTCTGGTCC	TCTAAAGTCA	GCTGGGGGAG	30180
GGGCTTGAG	GCCTCAGGTT	AGTCCTAGTT	AAACAGAGCT	AGCCTTTTCA	GACAACTGAT	30240
CTCCTTCAAA	AGACCCAACT	ACTGCCTTCC	GTTTCCCCGT	AAGTTCAGAT	GTAAACCTGT	30300
CCAGACCTTC	AAAAGTCCTA	CTGCCTCTGA	GCTTGAGCTT	TTTCAGTGTG	GGTAATGGGG	30360
AATTTTGGA	CTGAAATTAA	GTCTACACTT	AACAAAGGAA	GGAACCTTTC	ATCTACAAAT	30420
TCAGCCACCA	GCCAGCCTTT	CCGGTTTCCA	TCATTTCAAT	TGGATCATCT	AGACCAAGTT	30480
CTGGAATAAT	TGCTTAGGTC	TTCCCCACC	CCCACCCCA	CCCCACCCCT	GGCCTGGTAG	30540
ATCCCCCTCT	CCACATCCCT	GTTTTCTTGG	TTACTTCTCT	TCAGATTTAG	TTTTCCGTGA	30600
GGCAAGAGTG	GAGAAGGGAG	AGATGTACTA	GCCTGTGCTC	CTGTGTCACA	CTCTTGCTAC	30660
TCAGTTCCAC	TCTTAAAATT	TCTGGTCCCA	GAGGAATAGA	GATGACCTCA	CATGCAACCC	30720
TGCCTTGACT	ACTTTTCTAT	TGCTCTAAGG	AGGCAACATG	GCCACAGCAA	CTTGTAAGG	30780
CATTTAATTT	GGGGTTGACA	GTTTCTCAGA	GGTTGAATCC	ATGACCATCA	TGGTGGGAGC	30840
ATACCCGGAG	GCAGGCATGG	TGGACAGGCA	GTCGTGGGAT	GGCTCTGGAG	CTGTTGCAGA	30900
GCACTTATTT	GCTGATTGAA	AGCTCAAAGC	CTACCCCCAG	TGACACACCT	CCTCCAACAG	30960
GGCCACACCC	CCTAATCCTT	CTCAAACAGT	TCCACCAAGT	ATTCAAATAT	ATGAGCCTAT	31020
AGGGGCCATT	CTCATTCAAA	CCCCACCCCC	ACCCCGTGG	CCCTACTAAG	GGCATCAGAT	31080



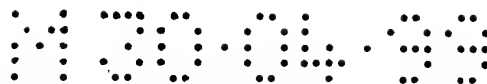
AGGGCCTATG	GAAAAGTTAT	AAACCCTCTC	ACCACCACTC	TGGGTTCCAG	CAACCCAAGG	31140
CCACCATTTT	CTACTCTTGC	TTAACCAACA	CCACCCAGGA	TCTCTCAGCC	TCAGCCTGGA	31200
ATGAGGGAAC	CCTCTTGTCT	CTTTTCATTC	AACTCCGTAT	TCTTCCTTCA	TTCCACCCAT	31260
GGATGGAAAG	ATTCACCCCC	TCCACTGTAG	AGTAACACAC	ACGTATGACA	AGCCACTTCA	31320
CTGCCCTGCA	TCTTACTTCT	GCTCTGAAGT	TCTGTCAGCC	AAAACGTATT	GAGCACTGAA	31380
GACTGTCACT	TGCTGCTTTG	TGTGGTGGTT	ACAAGTTAAG	GTCCGACTGT	AGCTGTCTGC	31440
TTGCTGGAGA	GACTGGGAAC	CAGTAGTTGC	TTAGCCCATG	GGGCTGGAGA	CCTCAGCAGT	31500
TCCAGTGTGG	TTCTGAGGAG	AACCCATTCC	AGCAGCAGCA	GAGGTAGCCA	CAGGATAGCT	31560
TGACTCACAA	GACTCATGAA	CTCAAGAAGA	GGAGAGATGA	ACTTGTAAGC	AGGGTATGTG	31620
AGCTCACACC	TGAGCGGTGA	AGGCAAGCAG	GTAAGAAGAG	CTTCCCCTCG	GACCTTCTGT	31680
CTGGGCCATC	TACACTCAGA	TGGGCCTCCC	ACTTCATTTA	CTAGAAGCAA	GCAAATCCCT	31740
CTCAGGCGTG	CTGAGGTAA	CCTAATCGGC	ATAACGCCTC	ATAGGTGTAC	CCAGAGCTTG	31800
TCCCGTGATA	CTAGATCCTG	TCAGGTTGAA	AATGTTAACC	ATCTCAAGGG	TCGTACACAT	31860
TCCAAAAAGG	CACTGTGTTG	GCTATTCTTG	GTTGTCAACT	TGACTACATC	TGGAATTAAC	31920
TAAAACCCAA	GTGACTGAGT	ATGCCTGGGA	GGGAGATTTT	CTTAAGTCAT	TTGAAGTGGG	31980
AAGACCCACT	TTAATCCAG	AACTTCTAAG	GTGGGCAGAT	TCACCTTTAA	TCAGCCTATT	32040
TCAATGACAT	GGAGGATGGA	AGTTTGTCT	CTTTCCTGCT	TAGCCCTTGT	TGGCAAGTCC	32100
ATCACTTCAC	TGAACCAAAG	CCTGTAAGGC	ATTCTTCCTT	TGTTTGTGTTG	GACAGGGTTT	32160
CCTGTAGCCC	TGGCTATCCT	GGTATTCACT	CTGTAAACCA	GGCTGGCCTT	GAACCTCAGAG	32220
ATCCAAGTGT	CTCTGCTTCC	CAAGTGCTGG	GATCAAAGGT	CTGAACCACT	AATAAATTGT	32280
GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	GTGTGTGTGT	ACACATATAT	ATGAGAGGGA	32340
GTGAGAGAGA	GAGTCATTCT	GTAAATTCTG	TTCCTCTGAG	AACCCTGACT	AATAAAGCTG	32400
CAGACTGCTT	AGTATCCTTT	TTGTTCTCTT	TGGGGACACA	CACAAATGAG	TGAACGGACT	32460
ACAGTGGGCA	ACATTCTTCT	ATGTCTGGTG	GCTGCCCTGG	GGCTGTTTAG	TCCACCCTTG	32520
TGTGAGGACT	CTTTTGCTCT	CAAGTGCTGG	CATCTGACCT	GTGCCCTTTT	AAATCTGTTG	32580
CTAATTTTGT	CTCTGGGGTT	CCAAGTAGAG	ACTTTTCAGT	GATCTTTCCT	CATGATGAAA	32640
ATGGGTGATC	TGTTATTGGA	AGTCCTTGGC	CTAAGCAAGC	TCTGATTTAA	TCTAACTATA	32700
TCATGTGCTC	TTCTAATCTA	TTGCTCCGGG	TCCCTGAGCA	TTGCTGTACT	CATTCATGGG	32760
TCATTTTGTC	ATTAATCTGG	CTCAATCCAT	GTTCACAATG	ATGATTTGAT	AAAGGCTGAA	32820



AATGTGAAGT	GGATGGTAAC	AGTTCTGTGC	CCTGGATTCC	AACAAAGAGA	TGCATGCTCC	32880
TCCAGCCCAC	TCTGGGTGAC	TCTAGGGGAC	GGAGACAAGG	GTCTTACAGA	GATGTCAGAG	32940
TATCTGACTC	CTTGACAGCT	AGTGGCCTCA	CAGGGAGACT	CATCAGGGGT	CAATGCTCTT	33000
TCTGGTAAGA	TGAACTCCAG	CTCACCCTGC	ATCTTGATCT	GTCCACACTG	CTTGGTGTTG	33060
AGACTTCCTG	TAGCCATGTA	AAGTGGGACA	TCTGGCCTAC	TGGTGATTCT	CTAAGAAGGA	33120
ATTTCCACCA	AGCAGGACAC	CTGAACACTT	TCTTAACATT	GA CTCTTACT	TTGGCTACCA	33180
AAAGAAGCCT	TTGAGCCCTA	TGTGGTAGCA	CAGACCTGCA	ATCCCAGTAC	TCAGGAGGTA	33240
GATGAGGTGG	ATCTGGAGTT	CTAGGTCATC	CTTGGTTGCA	TAGCAAGTTT	ATATTTGAGC	33300
TTGGCCTTGG	CTGCATGAAA	CCCTTGTCTT	CCAGGAGACA	AAAACAAAAA	CAGGCAAATT	33360
TCCCTTAAGA	AGCTCACACT	CCGCCTATCC	ACTGTGCTTG	CCTTCTTCCC	AATCACTATG	33420
GCCTCCTCTC	CTCCATTAAC	GCCCATGCTT	AAAGGGTCTT	CTAAAAATGT	CTTTTAGTAA	33480
ACTCCAATTC	TACTACATTT	AAAGAAGGGG	GAAGGTGAGC	CCCACATGCT	ACACCCACACA	33540
GTTCCAGGGT	GCTAGGCTTC	CGGCTGGGGG	CTGCCTCTTG	GTACTGCCTT	GCCCTGGAAT	33600
GTCAGTTCAG	CTAAAGGCCT	CACACAAAAG	ATGAAAGCCC	TGAGTCCTCT	TACTGCTTCT	33660
TAGCACACAA	GCAGTTTCCT	TCACTCCCCT	AGGTCTTAGC	AGGCCTTCAT	CTTCAAGGGT	33720
TCTCTTTCCC	TCTATTCTGC	CTTCTCTGTC	TCTCTCTCTC	TCTCTCTCTC	TCTCCCTCCC	33780
TCCCTCCCTC	CCTCCCTTCC	TCCCTCTCTC	TCTCCCTCTC	TCTCCCTCCC	TCTCTCTCTC	33840
TCTCTCTCCC	TCCCTTCCTC	CCTCCCTCCC	TCCCTTCCTT	TCTTTCCTTT	CATTTTCTTT	33900
CCCTTTTTGT	CCCTTCATGA	GAAAAAGCAT	ATTTGTAAAT	CCCAATTTAA	AATATAAATA	33960
AACGAAAACA	GTAAGTCTCA	ACCAAATGAG	GCCTAAATCA	GCCCTGGAAG	ATTAGTACCT	34020
GTTTCTACTC	AAGTTAATAA	TTTACTCTGT	GTCCCTCTGT	GCATGCTTGG	CTTCAACAGA	34080
GGATCTTTAA	CATGGGATGC	AACTTCGCCA	GAGAGCTTCA	GTTCTCAGGA	GGCATGTGGA	34140
CATCGTGGAG	GTTGAGGAGG	GGCAGATGGA	TGCTGGGAAG	CAAATGGAAA	GCCTGAGGTT	34200
CCAAGTCAAA	TCTGTGACTC	ACGCAGTAAG	GAGGTTTGAG	CTGGGGCTGC	CCAAGGGAGG	34260
AGGGCTACTA	CAGGCAATGA	TTAAGATTTA	TGTATTTATT	TTATGTATGA	GTACACTGTC	34320
GTTGTATAGG	TGGTTGTGAG	CCTTCATGTG	GTTGTTGGGA	ATTGAATTTA	GGACCTCGGC	34380
TCACTCTGAT	CAACCCCGCT	CGTTCCAGCC	CAAAGATTTA	TTTATTATTA	TACATAAGTA	34440
CACTGTAGCT	GA CTTCAGAC	ACACCAGAAG	AGGGCATCAG	ATCTCATTAC	GGGTGGTTAT	34500
GAACCACTT	GTGGCTGCTG	GGATTTGAAC	TCAGGACCTT	CTGAAGAGAA	GTCCGTGCTC	34560



TTACCCACTG	AGCCATCTCA	CCACCCCCTT	AAATTGTTAT	TTTTAAAACT	ATATGAAATA	34620
AACTTTACCA	TCTAAATGGG	GAGGGGTGAC	CAGTCTCCGC	ACATAGGAGG	TATAAGGGCA	34680
GGAAGATCAG	ATCTTAAAGG	TCAGCCTACA	TGAGACCCTG	TCTCATAAAA	ACCAAGTAAT	34740
TAATAATAGC	AATTAATAAT	TAATAATAAT	AGGACAGCAG	TAGCACTATT	TGGTTGCTGG	34800
GGATACAGCT	CTAGTAGAAC	ACTTAGCCAA	AGGGTCCTAA	ATTCAATGTT	GAGGACAGCC	34860
AAAAATAAAA	TAAAAAGTTC	CATGTTGTTC	CCCCACACAC	ACTTTTTTTT	TTTTTTGAAT	34920
GACTCTCACT	ATGTAGCCCT	GCCTGGTCTG	CAATGTACTA	TGTAGCCTAG	GCTAGCCTCA	34980
TACTCAAAAG	AGGGCTAGCC	TGCCACTACC	TCTGCCTCTA	GAGTACTAGA	ATTATCAGCA	35040
TGCTCAGGCA	CACTGGGTCT	TGTTTGTTTT	TTTGAGACAA	GATCTCATGA	ATCCCCCACT	35100
GGCCTCAGAT	TCTCCATGTA	GTCAACGATA	ATCTTGAATT	TATACTGGAA	AATGGTAGCA	35160
ATCTGGAGAG	TAACAAGACA	GGAGCTGACT	GTGTGTATGT	AGCCCAGGAT	GACCTTGAAG	35220
CCTGCCTTGG	CCTACAGAGC	GCTGGGACTA	TAGGGGTATC	CCACTGTGCT	TGCCTGCCTC	35280
TATGTAAAGG	TGGAACGAAT	TTCCCCTGTG	CCTGTGGACC	ACGTTTCTCT	GACCCACTCA	35340
TCCACCAAGT	GGCGTTTGGC	TTGACCCAC	ATCTCTTGGC	CACTGGGGAT	GATCTGAACC	35400
CAGTGCATT	TTCTCAAAAT	ACACTGAGGT	GGGATCATTG	GATCACAGAC	GTTCTTAGAG	35460
CCTAGCCTAC	CCCCTGGGGC	TACAGGAAGC	TCACAGTTTC	TGTTGGTTGA	TTGGTTGGTT	35520
TGCCCCCTCC	CAAACCCCTG	CCACCTCCCC	CCAACCTGGG	TTTCTCTCTG	TGGCTCTCTT	35580
GATGTCTTCA	AACTCACTCT	GTAAACCAGG	CTGACCCTGA	CCTCAGAGCT	CTGCCTGTCT	35640
CTGCCTCCCT	AGTGTTGGGA	TTAAAGACAT	GTACCATCGG	CTATACCTAC	AGACGTGCTC	35700
AAGGTATGTA	CAGAGCACTC	ACCCTGGCAT	CCCTTCACCT	GCCTAAGAGA	CTAAGGATCA	35760
GAAGTAAACC	CTACCTGCTT	CTCTGGAAGA	TTCAGGTTTT	CCTCAGGGTA	CTGCAGCCTC	35820
TCAACCTAGC	ATGGTCTGGG	CCTTATCCTT	ACGAATGTAC	ACTCAAACAC	AAAGACAAGG	35880
CTCTCCCAGC	CTGCCCTAAT	AACTTTTTTC	ACCAAACAGG	TCATGAGTCA	ATGGTGCCCC	35940
GATATTGTCT	AGGCAATAGT	CATTCTGGGA	CTACAGGCCT	TGGTACCCAA	CATGACTCCC	36000
TCAAAGCCAA	GATTGTGAGC	ATGTCACTGA	GGCCACTCTG	TGAGCTTGTT	TCCATGTCAA	36060
CGGAGCTCAT	GATGTCAGAA	GGCTGAATCC	AGACCCTGCA	CCCAGGCTGT	GTGTTTCCAG	36120
CTCCACCCCA	GAGCATATCC	CAGTCCAGCT	GGCTCTTTGG	AACCATTAAA	GAGTGATAGG	36180
TGCTGACTAT	GTGTGCAGAG	AGTGATCCTA	GCAGCACAGG	ACACAAATCC	TCACCCTGGG	36240
GAAAGCAGCC	TTCAACCTCT	CACCCTTAAG	GGGAAGGGCA	ACCATGGAAC	AGCATCTGTC	36300



AGCCCTCCCT CACAACCCCC CAGGCTGGCC TAGCCACACC CTGCCACTTC TATCCAGGCA 36360
 GCAGGGCTTC CTTTCCAGAG CAGGGGGGGT GGGGTCAGGG AGGAGCCTGG GGATTAGGGA 36420
 GGGACACTGA GTTCTTCAAG CAAGAACTGT TCCCCATCTA AGGCCATCCC CTCCTCCAGC 36480
 CCCAGCTATG CAGGGAGCCT GGCTGCTGCT GCTGCTGGGC CTCAGGCTTC AGCTGTCCTT 36540
 TGGTGTCAAT CCAGGTAAGG AGGCTCCCCCT AACTGCTTGT CCCCCTCAC AAGCACAGCC 36600
 TTCCACTGAC ACCTGCCTCC GGTCTCCCCC TTGGCCAGTG GAGGAGAAGA ACTCGGCCTT 36660
 CTGGAATCAA AAGGCGAAGA AGGCCCTGGA TGTTGCCAAA AAGCTGCAGC CCATTGAGC 36720
 ATCAGCCAGG AACCTCATCA TCTTCTGGG AGACAGTGAG TGTGTGAGCA CGGCCTGGCC 36780
 ACCCTGGGGC CCCCTGAGCT CCAGGCATCC ATTGATGTGT CCAGGAAAGC CTGGTGTTC 36840
 GATCGAACCA GATTCTGTTT TTGTAGGGTT GGGGGTGCC ACGGTGACAG CCACCAGGAT 36900
 C 36901

(2) ANGABEN ZU SEQ ID NO: 7:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 13330 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Genom-DNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Homo sapiens

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 7:

GATCCACCGC GCCTTCCCCG AGCAGGCCCT GCTGCGCCGG CACCCCCCGC CCCAAACAAG 60
 GATGCTCAGT GACCTGGTGG AATTCTGCGA CCAGATGGGG CTGCCCCTGG ACTTCAGCTC 120
 CGCAGGAGCC CTCAATGTGA GTGGTGGGCA GGATTCGGGG GAGGCCCTGC TTGGGGGAAA 180
 GAAGAGAAAG ACCTGGAAGG TGGGGTGGTC CAGCGGCCTC TGCTTCCCCC CAGAGTCCCT 240
 CCCCTTCAGC CAGGTCTCTC CTGTAGGGAA GGAGGCCCTG GGAGAAAGGG CCCCTCTGAG 300
 TCACAGGGGC CCTGACAGTG GGACCTGCCC CTTACCAGG ACTGTGCCAA GCGGGGGGAC 360
 CCTGGAGGCC TAGCAGAGGG CAGGGGTCCT GTGGCCAGAA AGGGCTGGTC TTGGGCCCAG 420

AGGCTTTCAG AGTCGGGGCT GGAATTGTAG GAATCCCGGG AATGTTCTCG GTGGGTACTT 490
 TCAGGTGCTC CCTGCCTGGG GCAAAGCTAA GAAACCCAGG GCCTTGGCTG TGGTCCTGGA 540
 GGAGGGAGAC ATCTCACCCA GGCCCAACCC TGGGAGGGGA AGGCAGGTGC CCCAGGCCAG 600
 AGAGCTGGAG CCCAGTGAGT CCAGGCCAGC CAGCAAAAAC ATGGAAGTGT GGGCCACAGG 660
 GTGTGGGCGG CTGCCCCCTC TCCCCACCCA TCCCCTCTGA GCAGGGCTGA GCCCCACAGG 720
 CAACTCCTCC CCCCAGAGCC GGGCATGAGG TGCTCAGCGG ATGACAGGGC CCAGAGTCTC 780
 TGCCCGAGCT GGACCACACG TCACATAGGT TTCTGGGATT TGCTTCTAGA AAAGCCTGAC 840
 CCAAACATTT GGAGATGACA AGTACTCACT GGCCCGCAAG GAGGTGCTCA CCAACATGTG 900
 CTCCCGGCC ATGCAGGTAA GGAGGGCCCA GCCCCGGCCT CCCCTGCTCC CAGGAGCACA 960
 CTAGCCCCAG ACCTGTGACC TCCACGTGCA AGCACAGGCC CCCACCGTTC CTGCCTGCTC 1020
 TGGACATGGC TGGGTGGACG GGGGCTGCTC CTCCTCTGCC AGAGGGTGGG AGAGGAGGCC 1080
 GACCCAGGC AGCACCTAGG AGGGGGCACC CTGAGCCTCT TGAGTTTGAG CCGCTGTCTC 1140
 CTGCTCACAC TCGCTCAAGG ACAGAGTGCC CTGGAGCTGA GGGGCTACTG AGACCTCCTG 1200
 TCAGGCTGGG GTCCTGGAGG AGAGACAGGG TCCCATGTGG TTTCTGTCC CAGGGAACAC 1260
 TCCGCAGCCT CCATCCCCAC ATGTGGAGTC CAGAACTAGC TGTGAGCCTC TGGCCAGTGT 1320
 GGGAAAGAAG CGGACTTGGC CGGGGGCCTA GGCCTGGGCC TGCAGGGAGG TGGCAGCCTG 1380
 TGGGGTGGAC AGCTGGGCTT GCTCTGGGAT GCCTGTCACA GCGCCCCAGG CTGAGCTTCC 1440
 CCCATGCAGG GCCCCAGCAT CQTGGGACCA GGACCCAGA GGACCCTCGG GTCAGCGGGA 1500
 GCAGTGGATG CTGATGGGTC GGCTCTGGGT CCCACCCCGG CCCAGGGGCA GAGACAGGCT 1560
 GTATTTTAGG GGCTCGGTCA CTCGGCAGAT TCAATCTGTT CACAAGAACT GATGGCTTCA 1620
 GCTGACCTCA GTGGATTTAT TTTCTGACAC TTCAAGCTCT GCTGGGTTTG AAGCCATCAG 1680
 GGCCTGCTTG GGCCTGGTCA CCGTGACCTG CCCCCAGTCA CAAGTGTCTG CCCAGCCAAG 1740
 CACCTGTGGC ACCCACAGCG GAGAGGGGCT GGGCCGTGCC CACTGGGCTC TCTCTGTTCT 1800
 AACTGTCAGC GGCTCTAGGC CTGGCAGAGA AGGCACAGCA GCCCCTGAGT CCCAGAACTG 1860
 CCTCTGGCTC TGCCCTGCTG GGGCCCCCTC CATGTCCCTG CCTCTGACGC CATCACCTCC 1920
 AAGGAGGTAC AAGCCAAGCT GGAGCTCCAG AGATCGGAGC CGCTCCGAG TTAGCCAGAG 1980
 CCCGAAAAGC CTGCATTCTC CTGGCTCGCC TCCCAGGGAG CTCAGAGGTG CCCTTGCCCCG 2040
 GGAATCCGAT GGCAGAGAGT TACCAGGTCT GCGGTGCTCC TGTTCTCAG CCCCAGGAAC 2100
 TGGGGTGGGG ACAGGGCAGG GCAGCAGCAG AGAGCACAGA AAGGTGTGAG GGGGCACACA 2160

GTCCCCAGTG AGCATCTGCA TCAGGACACC AGGGCTGTCC GAGGGCTGTC CCAGGGATGG 2220
 CTGGGCCTGT GGGAAAGCCA TGGTCCCCAC CCATCCCACC CGACCCTGAG CCACCTCCAC 2280
 CAGCCAAGAG GGGCCAGGGC CCTTCATCAA CCTCACCAG GTCATCTGGG GAACTGGGCC 2340
 ACCACTGAGA ACAAAGCCCA GACATGTCTG GGAGTGGAGG CTGTGCCCAC CTCCCCAGA 2400
 GACTTGCCCC CGACTTAACC CAGGGCCCAG CAGGGGCTGG AAGGGAAGTG GAGTTAGGGA 2460
 GCGGAGCAGG TCACCATCAG CTGCGCCCTG GATTCCAGGG CCCGTGTGCA CAGAGTAACG 2520
 GGAGCCGGCT GTCTGTCTGG CCAAGGGCAC AGGAGGTGA GTGTGTACAG CAGCCAGGGA 2580
 GCAAGGGAGC CAGAGAGACA TACAGGCGTG ACCTTGGAAC TCTGCGAGGA ACCCGTTCAC 2640
 TCGCTCCCAG GCAGTAGCAC TGGCCCTGAC ACCCAGCCCT GAAAGCTCGG GGACTGCAGG 2700
 ACAAACAGCT TCAGGGGCTG TGGCCCCAGC TGGGACGGGC TATGCGCTGG TCCCTAGAGA 2760
 CTCTCGGTAT CTCCCCCTGC CCCAGTCCTG CCTCCTGCCC AGCACAAGGG CCTTTGGAAC 2820
 TCAGCCCTCT GTGTCTCAGC CCCCGGGAGG GTCAGGTGTC AGAGACGAGA AGGGCCGAGG 2880
 CTGGCAGGCC GGAAACTGCC TCCCTTGAAC GCTGTGGGGT GGAGTATTGG CGAGCACAGA 2940
 GGTGCCCCGG TGAAGCGTGG CTTCAGCTGG GCGGGATCAG TGCCAGAGGG GATGAGGACG 3000
 GCCCCGACCA AAGGTGGGCC TAGGCTGGAG AGGAAGCTCC AAGAGCCTGA GGCCCGTATT 3060
 GCACAGGGCA GGGGATCGCA TCCTGGGCTT TCTCTCCCTC CTCCCCTCT GGCCAGATGG 3120
 GAGGATGGAC GTTGCCCTCT TGAACAAAGA CCCACAGGCT CCTTGGCTTC TGCTTGTGTC 3180
 TCCAGCAGAC AGCGTCTGCA GCCCCTGGTC CAACAAAACC GCAGGCGGCC TCCTCCTCTT 3240
 CCTCCTCCTC ATTGTCTCTC TCGACCACCA CCACCTCCTC CTTCCACCAC CTCCTCCTTC 3300
 TCCTCCTCCG CTGTGCGCTC CTCCTCGTCC TCCTCCTCCT CCTCCTCGTC AGCAGTCGCA 3360
 GCCTCCTCGT CCTCCTCCTC CTCATCCGCA GTCGCTCCT CCTCCTCCTC TGCCTCCACC 3420
 TCTGCCATCG CCACGTCCTC CTCCTCCTCC CCCACCCCCC GCCGCTACCT TTCTTTCTTC 3480
 TTCCTTCTTC CTGGGCGAGA GTAGCAGCCC CGGCCCCATG CTGGGGAAGG GTAGGCCAGA 3540
 GACTCTTCCC TCCTGGTGGT GCTCAGCAGT GACTCAGCAG GGACTGGACT TCGGAGGCTC 3600
 AGCTCGTGCC CCCTACCCTG ACAGCATCCT GGGGGTTCTT GGCTCCCTGG TCCTCAGCAG 3660
 GGTGGGCTTG TCCAGGCCAT TCTCAGTGCT GCCACCTTGA GGGCATCTGG GAGGCCCAGG 3720
 CAGGCCAGAT TTGTCTCCTG GAAAGGACAT GGGTACCCCT GGGCTCTGCC CAGCCTCCTG 3780
 GCCTCCCCCT GGGGCCCTT GTGCAGCAAG GGCCCTGGCC CCAGTCCTCC CTGGCGTCAC 3840
 TCAGCAACCA GCAGCCCATT AGGTCTGTCC ACACATCGCT GCCGACGGTG AGGCTGTGGG 3900

TGGTGCCAGC CTTCCAGGCC TGGCTGGGCA GCTCTGGGCT TGTCAAGGCTC TGACCCATCC 3960
 CGTCCCGCAG ATGGCACTGT ACTTCTGCTC GGGGCTGCTG CAGGACCCAG CGCAGTTCCG 4020
 GCACTACGCG CTCAATGTGC CCCTGTACAC ACACTTCACC TCGCCCATCC GCCGCTTTGC 4080
 CGACGTCCTG GTGCACCGCC TCCTGGCTGC CGCGTTAGGT GAGGGGTGCA GTCGGGGTCA 4140
 GGGCAGACCT GGGCCAGCTC AGGGCTGCCC ACCCCCACAG TGGGTGCTCA GTGGCCCAAG 4200
 ACCATTCTGC CGTGACAGCG GAGGTCCAAG GGTCTGGGCGA CCCAAGTGCA GGGGAGCCTG 4260
 GCCTGGAAAC TCTCCCTACG GGCCGGTGCT GCAGAAGCTG CATGGAGCCC ACAGCCAGCC 4320
 CTGGACACAG CCGGGAGGAG GCGCTGACC TCGAAGGGCC GCTTTCTGCT GCCCTGGGAG 4380
 CTGGGTGCTT GGGGTCCTAA TCTGTCGGCG GGGGTGCAGC GCCATGCAGC CCATCCCCCA 4440
 GCCATAGCTC TTCCAGCCC CCCAGGCTCC CACTCTCATG CCTCACCCCC TCTTCCAGG 4500
 CTATAGGGAG CGACTAGACA TGGCGCCGA TACCCTGCAG AAACAGGCGG ACCACTGTAA 4560
 CGACCGCCGC ATGGCGTCCA AGCGCGTGCA GGAGCTCAGT ACCAGTCTCT TCTTTGCTGT 4620
 TCTGGTCAAG GTGAGCCCTC CAGCCTGGTG CCCCTCACCT CCCTCTGGCT CCCGACCTC 4680
 CTGGGCACCT GCTCACCAGG AGGCCTCGAG GAGCCCAGGG CAGTGCCAGG AGGTGCCATG 4740
 GCTGCAGCAC TGTCCCTGCA GGAGAGTGGC CCCCTGGAGT CAGAAGCCAT GGTGATGGGC 4800
 ATCCTGAAGC AAGCCTTCGA CGTGCTGGTG CTGCGCTACG GCGTGCAGAA GCGCATCTAC 4860
 TGCAACGTGA GTGCCCTGGG AGAGCCCGGG GGCGGGCAGG GCAGCCCAAG CCATCCCGCA 4920
 CTGGAGGGGC ACAGGCTGTG ATGGGTCACA CTCCACCCCT CGCTCCCCCA GCCCTAGCAC 4980
 AAAGCCCACC TGATGGGCCT TGCTGAGACG CCCAGCTCTC CCACCTGGGA TGGTGGCTCC 5040
 AGGCCCAGGG TCAGGCCTGG CCCCCTTCCC CAAGGACCCA GGAACCAGAG AGCAGGCCCC 5100
 TCCATGGCCA GTACAGCTCG GCAGGGTGTG CAGGCTTTGG GGAAGTGTGTT TATAGGAACG 5160
 TGAAGGAATG AAAGGCCAGC GAATGGTCCG TGGCCGCTTT GGAAACTGTG TCCCCTGAAG 5220
 ACAAGGAAGA GAGCTGTCCC TGGCTCGGCT CCTGCCCTGA GTGACTGTTG ACTCACAGTT 5280
 CTCTCTCCAA GGGGACATGG GCCTGTCCTA ATGCTGCCTT AGGGGCTTGG CTCCAGCTGG 5340
 CCCTGGGGTC TGCAGGTCAC CACCTGCCTC TGTGCCTGGC TTTGAATTTT CTAACATCCA 5400
 GAGTGCCCTG GGAGTACAGT GTCCAGCCCG TTGTGTGCAG TAAACGTGGT GTTCATAACC 5460
 GGGAGCTGGG CAGAAGAGGA ACGACAGAGT CCCCCTGCGG ACCCTGGGGG CTCTGTATCC 5520
 TGAAGTTCAA GCCTAGCTCA CCCTGCTGTG GGCCAGCCC TGCTGCACT GACAGATGGC 5580
 ACCAGCAGGG GGCGCAGCGC TCCGCCGCCA CAGTTCTCTG TCCCCACCTC AGTGCAGTCA 5640

GCCCTGGACC CCCACCACT TGCCCCCAT AGCACACAGA GCCACGGGGG TTCCAGGCC 5700
 CCACCCCTGG CCCTTGGTCA CTCTCACCTG CTGCCTCAGC TGAAGGTGGC CTGGCAGGGC 5760
 CTCCCTGAAG CTCCCTCCAG CCAGGCAAGG GTGGGCCAGG GCCGAGGGCT GAGGGCCGCC 5820
 TCCAAGCATT GAAGCCCTCC AGGGTGGAAG GGCAGGCAGC AGCATCCAGA GCTGAGGCCT 5880
 GAGGCTTGGT GTTTGCACTC CAGGCACTGG CCCTGCGGTC CCACCACTTC CAGAAGGTGG 5940
 GCAAGAAGCC GGAAGTCACG CTGGTCTGGG AGCCTGAGGA CATGGAGCAG GAGCCAGCAC 6000
 AGCAGGTCAG AACCCCTCTG TGTCCCAGCC CCCTAAGTCC TGATGACCCC TCTCCTGCCT 6060
 CCTGCGGTGC CCCTCATTCC TTCATCTGTG TCCCCTGGGC TCCCCAGCA CTGCAGCCTC 6120
 CCGGTGGGG TTTTAGGGCC CTCCCAGCTC ACCCAGACCC CCTCCTGTGG GTCCTGCTTT 6180
 CTGGCACCAC CTTCCCTTCC TTGGGGGCAA CCACAGTGA GAGAGGAGGG GCTCTGCCTG 6240
 TCCCGCTAAT GCAGGGGTGC TGGCCTTCTA GGGTCCTTTA GAGAACCTGA TGAAAGCTAT 6300
 GAGTTTACAC CCAAGAAATT GTCTGGAACC GTTTTCACCA ACAGTGTGCC CTGAACGCGG 6360
 ACCCAGGCCC TCAGGTTGTG TTTCATAAGC CTGGGAGCG CTCAGGATGC ATCTGACTCC 6420
 CCAACTCTGC CCTGACCCAG GGCATTCTTC CTGGAGGGGG CCCCATTAC AGACAGGCGA 6480
 GCAGAGGCTT CCAGAGGCCG AAGGAGGGGC CAGGGGTCCT GCTGCAGGGA TGGAGGCAGA 6540
 GCTGCGCCTC GACATCAGGC CCTGCCATCC TTGTCCCCTC ACGGCTGGGC TCTGCACAGG 6600
 TCATCACCAT CTTCAGCCTG GTGGAGGTGG TCCTGCAGGC AGAGTCCACA GCCCTCAAGT 6660
 ACAGCGCCAT CCTGAAGCGG CAGGCACCC AGGGCCACCT GGGCCCTGAG AAGGAGGAGG 6720
 AGGAGTCTGA CGGTGAGCCC GAGGACTCAA GCACCAGCTG AGCTCCACCA GCCGCCTGCC 6780
 CCGCCTGCCC CGCCTGCCTG TCCCGCCACA CTGGCTTTAG GACCTGTTGA CACGAGGGG 6840
 GGTTTTTAAT TTGGTTTTTA ACAACTCAGG GGTTTGT TATTTTTATT TAATTTTTGC 6900
 AGCTCAACTT TTAAACAAAC TGCAGGGGAG AGGGTGGGGC TGGAAGGAAG GCTGAGGCCT 6960
 GGTCAGCAGT GACCCCAGCA GAGCAGGCCC CAGTCCTCCT GGGAGGCTGG CCCCCCTTTT 7020
 TTCTGGGCCC TACTGCCCTC CTCTGCCAG GAAATGGGGG GGTTTCAGCA ACTCAGTGTC 7080
 ACAGAATAAA ATCAAGTGTG GAGTGCCATC TGGTGTGTAG GGCGCCTCTG GGAAGCCTGG 7140
 GCAGCAGAAT GCCCCTTGCA CCCAGGGCAA GGGACCCAGT TCAGGCTTCA CCCCTCGCTG 7200
 CTGAGCCGAT GTCAACACCT GGAACCTTCC TGTCAGTTCC AACACGATTC AGAGCTGGCT 7260
 GCCTGGCAGA TGATTGATAC TGGAGTCTCA TTCTGCCTGA TTAAAAATGG AATTAGTATG 7320
 CAACACTGAG AGCGCCCCCA TCACCCTGAC GAATGTGACT GTGTCTGACG AATGTGACTG 7380

TGTCCAACCC TGCCCCCACT TCCTCTCTGC ACCAGCTCCG GAGGCGCTGG TGGGAGTCAT 7440
 GGGTCCTGTG ATACCCCCCTC CCCTCAGTTC CTCAAGCAGC ACTCTGTGAG GTCCTGTGCC 7500
 CAGCTCTGGT GTGAGTGGGT GCCCCGGCAG CACCAAGGGA GCCTGGACAG AGGAGCCGGC 7560
 CTGGGCCTGG GGGAGGGGAG GAGGGCCCTC CAGTGCCTTC CAAACCAGGA GGGGAAACTG 7620
 GCTGCTGGTG ACACAGCCTG GGTGACACGG ATCCACCTG CCTCAGTCCC GAGCAGAGCT 7680
 GGCTGGCCAC TGGGCAGTCC CTTCCCCAGC CAGCCTGACC CCAGCCTGTA CTCCTTCCCC 7740
 CTCCGTGGGG GAAGCTCCGT GGCTTGGCGT CCCCAGAGAG TGCCAGAAAC TAGGATGAAA 7800
 GCCATGGTGA GCACGGCCTC TGTTCCTTG CACCATTTC TGGGGTGTCC GGATTAACAA 7860
 GCTCATTTGA TCTGGTTACA GTGAATTTTC TTCAAAGAAA CACTCAATAG GGTCCCTTGT 7920
 CAGAGTGCCT CGCAGCGACA GTGACTGGGT ACTGCTGCCT TTGTCCTGCC ACCGTCAGAC 7980
 GGGGCTGGCT ATGGGAGGCA ACCAAAGACA TCCCGCACCT GCCCTGGGAG CCTTTCCCTC 8040
 CTCCAGGGCT CAGCCACCTC AGGCGGCCTT CCGTCTGTGT GTCCTGCCAC CCCCAGATG 8100
 TCCCAGAGGC CACGGTCACC CCATCTGTTC CTGTCCCCAG AACCTTCTCC TGGAGCCAAG 8160
 TATCTGCAGG GACAGACAGG CGAGCGTCTG GGGGTTTGGT GTTGGGGTGG AGAAGGCTGT 8220
 GGGGTGCTGC CCCAGCCCAG GCAGCCTGAC TGTGAGAGCC CCAAACAGGA GAGCCCCAAA 8280
 CAGGAAGGAC CAGGGCCCTT CCCCTCCCCT CCATGCTGCC CACCCTCTGA GGAGCAGTGG 8340
 CCAAGTTCCT CTCTGGGCTT CTCGGGCCAG GCTGACCCTG TCCCCAGGG CCTCCACGA 8400
 AGCATGGGAG CTGTTCCCTC AÇAGGCAGCA CAGACCCGGA CGGACACCTG TCCCTATGTC 8460
 CCAGCGCCCC CAGGCCCCAG TGAGGAGTAG CCAGGGGGGT GAACAAGGGG GTTCCTGCTG 8520
 CCTGGGCTTG TTTGGGAAGC AGATGCTGGG CTCAGAGTTT CTTCAGAGAG CCTCACCTTC 8580
 CGTGCTGGCC CCAGAGCATG GCGGGTCCCT GGAGCTGTGG AGGCCATGGC AGCCCCAGCC 8640
 CACCCACCC CATCTGGGGA AGTGGAAACC GTATCCACGA GGGTCAGGTC AGGTCTCTGC 8700
 CTCCAGTGAC CTGGCAAGGT TGTGCCCAGC CAGGACCTGG GCTCAGGCC AGGCAGCCGC 8760
 CACACCCTAC CCAGAGCTCA GAGAAGGCAG CCCAGCCTTC TCCCCACACC AGTCACACCG 8820
 AGCCCCGCGT CTGCATTAC TCCTTTAAGG AACATGGTTG ACTGAATCCG GTGCCGCGCA 8880
 TTCACAGGAT GGCTCTCCAT GGGTCCACTG GGGCCCAGCC TCTTATGTGG CCCCTCGCTA 8940
 AAAGGACTCA ACAGAAAGAG TGACCAGGCA CCGACCCTCA TCTAAAGGAG GACTTGGCCA 9000
 TTCCCTGGGC TGTCCACAG CACCTGCCGG CCAGGGCCCG GGCACAGAGC GAGACTGTCT 9060
 TTTCTCAAG GAGACACCGT GGGGGAGGGA GGGAGAGGTA GACACCACCA ACCTCATTC 9120

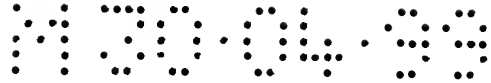
ATGACCAGGG CCTGGCGATG CTCAGAAGCC AGTGAGTGTG TCCCTGCCCT GAAGGGTCAG 9180
 TGCTGGCCCC CTGGACCTAG GGGGAAGATG GTGCAGGCAG TGGCCCAGCC TGAGGAAGGA 9240
 GCTGAAGCTC TCAAGAGTTT GCAGCCACCC TCCTGGGGAG AGACTGACGC CTCCCCAGTT 9300
 CCTGTTAGGA AGGACCTCAG GAAAGAACTG GAATTACACA GCCTGGGGTG GCAGCCTCCT 9360
 GGTCCCTGAG GAGGATGTCA GGCCGCAGAA GGGAGGAACG GGCATGAAGC TTGGGAAGCG 9420
 GGCGCCAGAG GAGGCGAGGC CTCTGCAGAA GCAGCACCAG AGGCCACTGC AGCGGCTCCA 9480
 CCACCCAGCA GCGCCGCCAG GAGGCAGGAA GTGGGAGGCC AGGCAGGAGG GGCTGTGATT 9540
 GCCCAGGTGC CAGGAGGAAG GGCTGAGAGG GGACAGTGCA GATGTCCAGA GAGGCCTGAC 9600
 AGGGACAGGC TGCGAAAGTC ACGGGTGGGG ATGGGCTTCC GCCAGAGTTG TGTGTGGCCT 9660
 GAGGACAGTG CAGCAAGGAG GCCCCATGGT GAGCACATGC AGCCGAAGTG ACAGGTTGGG 9720
 CTCCTTTGTG GGACAAGAGC CTCTCCAGGC CACTGCAGGG TGTTCAGAGA ACAAGGCCTA 9780
 CAAGGATCTG CTGTGCCTGC AGCTGGGCAG TAGAACACTG AGCATGCAGG GCCGGGGTGG 9840
 GAAGCAGGAA AGCCACATGG ACGAGAGAGC CGGGCCTGCC CAGCAGTGCC TTTTGGGAGC 9900
 GCAGGCAGGA TGGGATGTGC AGCTGTGACC TGCCCGGCAT AGAACTCCGT CTGGCTGGGG 9960
 AGAGGAGGTC TCTTCTAGCC AGAATGGACC AGGAGGTCCC GGGAGGACCT GGGAGGAAGT 10020
 GGATTGAGTT GGGCCTTAGA AGGAGAGCCA GGAACAGGCC AGGTCAGGGG AGCTGGAGCC 10080
 TGGCTAGGTA TGGAGAGAGC AGGGTACACT TGCTGCAACT GTGAGAAGAG CCAGGGGTGG 10140
 CCCTGGTGGC CTGGGCGCGT TTAGCTGTGC CTGGGGCCAG GCCTGACTGG CTGCAAGTCA 10200
 TTA CTATAGG CGGAAAGTGC AGAGTAGCGC GTCCTGCTG TCACTCCCTC CTCCAAGTCC 10260
 ACAAGAGGC AAGAAAGGGA GGATTTTAAG GCCTATCCAT ACCGCATGGC AGGTGAGAGC 10320
 AGAGGAGCAA ACAGCACTTT TGGATCCTGG AAAGCAGAAG GTGAGTGTCC CAGGCGTAGC 10380
 TGACCTGAGA AAGGCGACTC CAAAGCCAGC AGCAGCAACA GCTGGAAGT CCCCAGCCTG 10440
 CACCACGGGA CCCCAGCTC TGAGACTGAG AGCAGCTCTG GGGACCTCTG GGCTGGGGTG 10500
 AAGAGGGATG GCTGGAATCA TTGTTGCAAA CAATTCAGTA GGCAGGCAGC TCCCTAGATC 10560
 CCACCGTGGT CTGCAGAGGC CAGCACCTGT CCCGACCTCT TACTGGTCGG CCCTGGAGAG 10620
 CCATCTCCTA CAGAGGCAAA ATGAACGGTC TCTGGGCCAG GACCAGGCCT GTTCAGGGGG 10680
 ATGTGTGGCT AAGTGCATAA GGGATGCTGA GACTACAGCC CTCGTGCCCA GGCAGCGCTC 10740
 AGGGCATGGA TAGCCAGGCC CTCCCCATCC AGGCCAGAGA TGGGAAGACT CCATCCAATC 10800
 TCATTCCATG ACCAGGGACT GGCAAAGCTC TCAGTTCTCT CTCCATCCCA GCAGGAGACA 10860

AAGAACCCAA CCTCAGAGAT TCCTCAACTC GGAGACCCAG CCAGGCCACC CTCCAGAGCA 10920
 TCTCAGTCTG CAAGCCCCTT GGTGTGCTCA GAGCTTCCAG TCACACTGCT CATGCCTATC 10980
 CGTGACACAGC CAGGGATTGC CCTTCGTGGA GGAAAACTTC ATGAAACAAA AAACAAGCTC 11040
 CGTGGGGAAC ACAGACCATA GAGGAAAAAG AAAGCTGTAG AAAAAGAAAT GATGAATGCC 11100
 TTCCTGGAGG TGAGAAAGCC ATCGTGAAAC GAGAGGAGGT TGCTCCAAA AGTTCCTAGA 11160
 GAGCAAAACA AGGGCCCTTG GAGGCACAAT GATTGCCACC GTGGAGACAC ATTTAGCGC 11220
 CACTAGAGTA AAAACACTGC AGACAGGTGA GCTCTCAACA GATACATGTC CCTCGCCTTC 11280
 TCAGGAAAGA TGGGCAGTAA TGAGGGCAGA AGCCACAAAG AGGAAACCGT AGTGACAGGA 11340
 CCCAGGGTCC TTCAAGCTGC GGTGGGGCAA GCGCTCGGGA CAGTGGTGAG GGAGCAGCTC 11400
 AGCCCCAGGT GGTGCCTGGC AACCCGCCCC GGGACGTCCC ACCCAGGGCA GCAGTAGAGT 11460
 GACATGGATA GAAAGCTGAA TTCCCCAGAA GAGCCTGGAG GACATTGAAG TACTTCGCAT 11520
 AGAGCCTCGG GTTGGATTAG TAGTACATAC AGAATGATCC ACATGTGAAG ATAAGACCAT 11580
 GATTGGCTCC AGAGAAAACA GCAGTGCAAG CAAGAAGAGG TAGCTAGTCA CAGTTTACGA 11640
 TCTGGCAATA GCGTTTACAC AGTCATCACC ATAGAAATGC CGAGTCAGGA TCTAGTTTAC 11700
 TGCAGAACTC TATCAGGAGG ACTGGAAGAT GGGGACGCTG TCCACATGCA GGAATGCAG 11760
 TTGGTGAAAT GGAAGCTAAA TGCTCATTTT CCTCAGTGGG AAGCTGTGGC TTGAAGATGA 11820
 CTGTAAACTC TCTTTCCGCC TCTTCAATCT TGACAGGCCC CAGGGCTGCT AAGCTAATAT 11880
 GGCAGAAGGG AACTGTGCC AGTTGCAGGC CCAGGCCTTA AGAGACTGGC AGCTTCCCCT 11940
 CTCTGTCTCT GGAAACCTAC CTGCCCTTCT GTAAGGAAGC CCAAGCAGCT CTGGAGAAGC 12000
 CCTTATGGAG GGGCCCACTC TCAGCCCACA GCCAGCACCA GTTGGGCAGC CACGCAGACC 12060
 CCCAACCTGC AAGCCAGGCC CGCTGAGGCC TCAGTACACA CAGGCAGTCC CATCAGCCCT 12120
 GCCCAGATGG CAGTTTTGTG ATCAAAATAT AGACGATAGA TGATTGTTTT TTAAGGTTGT 12180
 TGGGGGTAGT TTGTCACACA ACGATAGATA ATAGAACATC AGTAGGCTGT GTGTGTGTGT 12240
 GTGTGTGTGT AGCATATATA TATACACATA TACATATATA CACATATACA TATACACATA 12300
 CACATATATA CACGTATACA TATATACACA TACATATATA CACATATATA CACATACATA 12360
 TATACACATA CACATATATA CACATACACA CACATACATA TACATATATA CACATATATG 12420
 CATATATACA CATATATACA TATATACACA TAGCTTCAAA TTCAGACATG AAGAAGTATC 12480
 TTATTTAGCA ACAGTGGTAA ATAGTAAAC ACCAAGAGAG AGGAAAGTGG TTGCCTCAGA 12540
 GATGGGAAAA TGCAAGGAGG GAGACGGAAC TGCTGTTTGT TTTAACAAAC CTTGTAGATC 12600

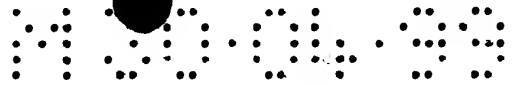
TGTTTGATAC TTAAACTAC ATTCACATAT AACTTGGACA AAGTAAAA CTGAAGTTGA 12650
AAAAAATGTA TTCATGCTAA TAGCACAGGA ATGATCCACA ATTGGATTCC AAGGCTTCTT 12720
GTACATTCAG CATAGGGTGT ATGAAAGAGT CCACTATTCT AGCAACAGAT AAAATTCCTA 12780
CTGACACGCA ACCTCAGGTT CCCACTCGTT TAGAAGGCTG CGTATGGTCT TCTACTTAAA 12840
GCCTCAAGTA GCAGTCATGG CAGTGACAAA TCCTCATTGC CTCCATAGAA CCTCTAGGCT 12900
CATGTGTGAG CCCAGGCTGG GCTGGGGCCC CTGGGAGCCC AGGGTGAGGG GCCAGTCCCT 12960
GGGCAGCTCC GTGAGCCAGG AGCAGCTGTG CCACCTGGGG AAGGGCTGCA CGGTCGATGG 13020
GTCTTTTCTG CAGAAGAGTG TGCCCCAGCC CTTGCTGGGC ACAGATCAAA GAGGTGTTCA 13080
TGGGTCGAAA TCACAGATT TCAAGGGCTGA TAGGAGTCAG AGTGGGGGGG CTGGGAGGGC 13140
TGAGGCAGGT TAAAGATTG AGAGGGGCTG CTGTGTCCAC AGCTGCATCA CACTGCTCTG 13200
CTGTCCCCTC CATGTTCCCC GGCCTGCCC CCTACCCTGG GGTCTTCTGG AAGTAACTGA 13260
AGGCCCCCTC AACCTGGCTC ATCATCAAAG CAGACTGTTG ACTAGCTGCA GGCAAATATG 13320
AAGAGGCTAT 13330

P a t e n t a n s p r ü c h e

1. Nucleinsäuremolekül umfassend eine Nucleinsäuresequenz ausgewählt aus der Gruppe bestehend aus
 - (a) Nucleinsäuresequenzen, die die in Seq ID No. 2 oder die in Seq ID No. 4 dargestellte Aminosäuresequenz codieren;
 - (b) Nucleinsäuresequenzen wie in Seq ID No. 1 oder Seq ID No. 3 dargestellt;
 - (c) Nucleinsäuresequenzen, deren komplementäre Sequenz mit den unter (a) oder (b) genannten Sequenzen hybridisiert; und
 - (d) Nucleinsäuresequenzen, die von den unter (c) genannten Sequenzen aufgrund der Degeneration des genetischen Codes abweichen,wobei das Nucleinsäuremolekül ein Protein codiert, dessen Inaktivierung in Tieren zu einer Verlängerung der Knochen mit Ausnahme der Schädelknochen führt.
2. Nucleinsäuremolekül nach Anspruch 1, welches genomische DNA ist.
3. Nucleinsäuremolekül nach Anspruch 1, welches ein cDNA-Molekül ist.
4. Nucleinsäuremolekül nach Anspruch 1, welches ein RNA-Molekül ist.
5. Vektor enthaltend ein Nucleinsäuremolekül nach einem der Ansprüche 1 bis 3.
6. Vektor nach Anspruch 5, wobei das Nucleinsäuremolekül verknüpft ist mit regulatorischen Elementen, die die Expression des Nucleinsäuremoleküls in prokaryontischen oder eukaryontischen Zellen gewährleisten.



7. Wirtszelle transformiert mit einem Nucleinsäuremolekül nach einem der Ansprüche 1 bis 4 oder einem Vektor nach Anspruch 5 oder 6.
8. Verfahren zur Herstellung eines Proteins, das von einem Nucleinsäuremolekül nach Anspruch 1 codiert wird, wobei eine Wirtszelle nach Anspruch 7 unter Bedingungen kultiviert wird, die die Expression des Proteins erlauben, und das Protein aus den Zellen und/oder dem Kulturmedium gewonnen wird.
9. Protein codiert durch ein Nucleinsäuremolekül nach Anspruch 1 oder erhältlich durch das Verfahren nach Anspruch 8.
10. Antikörper gegen das Protein nach Anspruch 9.
11. Nucleinsäuremolekül von mindestens 15 Nucleotiden Länge, das spezifisch mit einem Nucleinsäuremolekül nach Anspruch 1 hybridisiert.
12. Diagnostische Zusammensetzung enthaltend ein Nucleinsäuremolekül nach einem der Ansprüche 1 bis 4, einen Vektor nach Anspruch 5 oder 6, ein Protein nach Anspruch 9, einen Antikörper nach Anspruch 10 und/oder ein Nucleinsäuremolekül nach Anspruch 11.
13. Pharmazeutische Zusammensetzung enthaltend ein Nucleinsäuremolekül nach einem der Ansprüche 1 bis 4, einen Vektor nach Anspruch 5 oder 6, ein Protein nach Anspruch 9, einen Antikörper nach Anspruch 10 und/oder ein Nucleinsäuremolekül nach Anspruch 11 und gegebenenfalls einen pharmazeutisch verträglichen Träger.

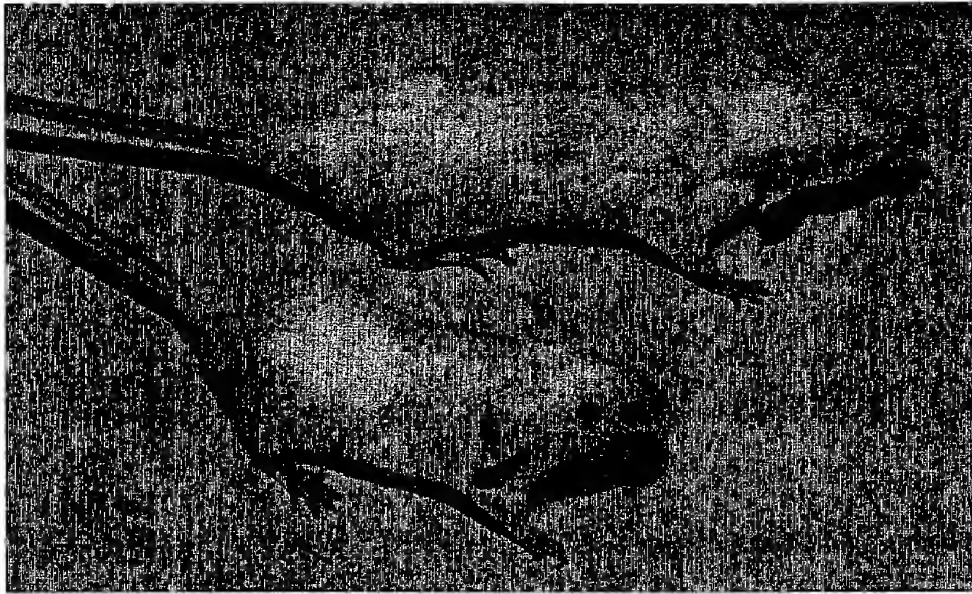


14. Verfahren zur Herstellung eines transgenen nicht-menschlichen Tieres, wobei ein Nucleinsäuremolekül nach Anspruch 1 oder ein Vektor nach Anspruch 5 oder 6 in eine Keimzelle, eine embryonale Zelle, eine Eizelle oder eine davon abgeleitete Zelle eingeführt wird und aus der so transformierten Zelle ein transgenes Tier erzeugt wird.
15. Transgenes nicht-menschliches Tier, das transformiert ist mit einem Nucleinsäuremolekül nach Anspruch 1 oder einem Vektor nach Anspruch 5 oder 6, oder das erhältlich ist durch ein Verfahren nach Anspruch 14.
16. Transgenes nicht-menschliches Tier, bei dem die Expression eines Proteins nach Anspruch 9 in den Zellen im Vergleich zu Zellen eines entsprechenden Wildtyp-Tieres verringert ist.
17. Transgenes nicht menschliches Tier nach Anspruch 16, wobei mindestens eine genomische Kopie eines Gens, das einem Nucleinsäuremolekül nach Anspruch 1 entspricht, inaktiviert ist.
18. Transgenes Tier nach einem der Ansprüche 15 bis 17, daß ein nicht-menschlicher Säuger ist.
19. Transgenes Tier nach Anspruch 18, das eine Maus ist.

Zusammenfassung

Nucleinsäuremoleküle, codierend Proteine, die die Knochenentwicklung beeinflussen

Beschrieben werden Nucleinsäuremoleküle codierend Proteine, die die Knochenentwicklung in Säugern beeinflussen, ebenso wie die codierten Proteine, Antikörper und pharmazeutische und diagnostische Zusammensetzungen. Ferner werden transgene Tiere beschrieben, die die beschriebenen Proteine exprimieren, und Tiere, bei denen das entsprechende Gen inaktiviert ist und die eine Verlängerung der Knochen aufweisen.



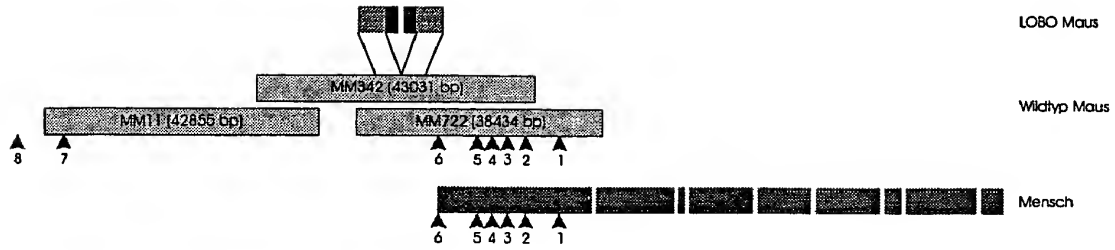
Figur 1

Figure 1 consists of five scatter plots arranged horizontally, each representing a different value of the parameter α : 0.05, 0.1, 0.2, 0.3, and 0.4. Each plot contains 1000 simulated data points. As α increases from 0.05 to 0.4, the distribution of points evolves from a single, somewhat irregular cluster on the left to a more structured ring-like shape on the right. The points are represented by small black dots.

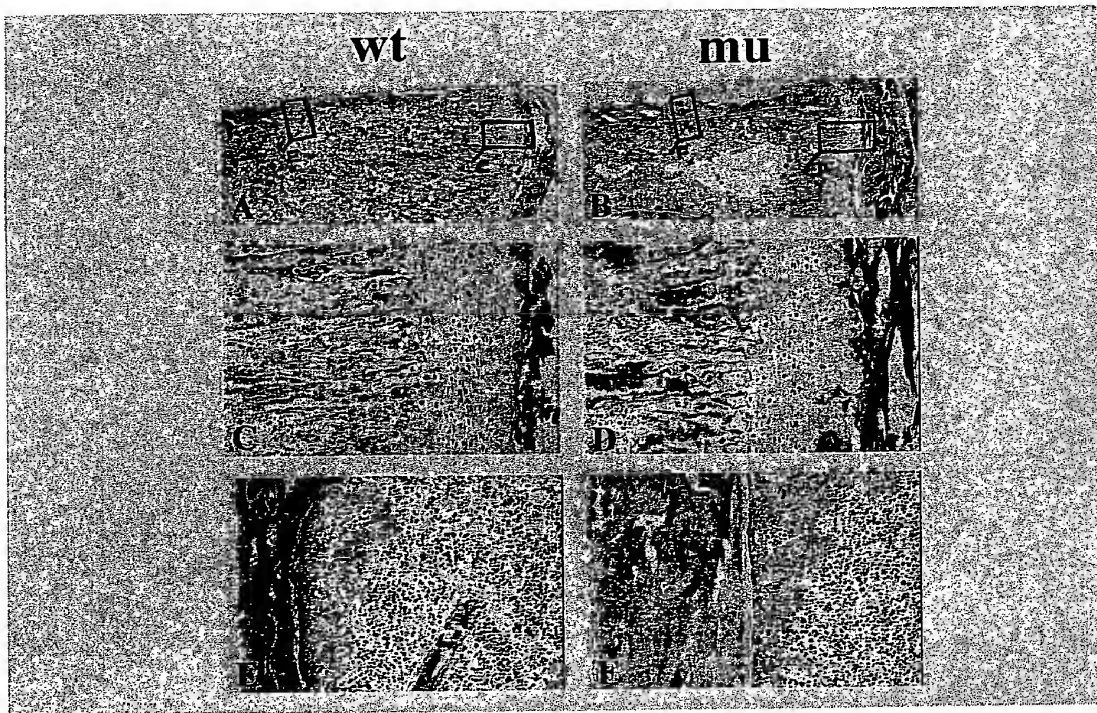
SC-Dis3
Spom1.p
Lactococ

Figur 2b

	840	850	860	
SC-Dis3	L V R I M S T R C M M A A Q Y F Y S - G A Y S Y P D F R H Y G			
Spom1.prot	L F E N M A V R S L N R A E Y F C T G D F G E K T D W H H Y A			
CE	V I S S L T I K P M Q Q A K Y F C T - F E M P L S F Y H H F A			
HS	- -			
MM	- -			
Lactococ	- -			
	870	880	890	
SC-Dis3	L A V D I Y T H F T S P I R R Y C D V V A H R Q L A G A I G Y			
Spom1.prot	L S F N H Y T H F T S P I R R Y P D I I V H R L L E R S L K N			
CE	L N V D H Y T H F T S P I R R Y P D V I V H R Q L A A A L G Y			
HS	L N V P L Y T H F T S P I R R F A D V L V H R L L A A A L G Y			
MM	L N V P L Y T H F T S P I R R F A D V I V H R L L A A A L G Y			
Lactococ	L A A T N Y T H F T S P I R R Y P D L I I H R L L H L Y P S D			
	900	910	920	930
SC-Dis3	E P L S L T H R D K N K M D M I C R N I N R K H R N A Q F A G			
Spom1.prot	T S - - - P G I D K K N C S L V A A H C N E K K E K S T T V Q			
CE	N E R - - S E R V P E E I Q E I C T R C N D T K L A S K E A S			
HS	R E R - - L D M A P D T L Q K Q A D H C N D R R M A S K R V Q			
MM	S E Q - - P D V E P D T L Q K Q A D H C N D R R M A S K R V Q			
Lactococ	H S N R T K D E W K E R L P E I A S H S S D M E H R A V V T E			
	940	950	960	
SC-Dis3	R A S I E Y Y V G Q V M R N - - - - - N E S T E T G Y V I			
Spom1.prot	E D S Q Q L F L S V Y I A E Y C K K H D K K S M P V Q A F A T			
CE	D E S A M L Y F G V F I H Q - - - - - T G P M K C Q A V V L			
HS	E L S T S L F F A V L V K E - - - - - S G P L E S E A M V M			
MM	E L S I G L F F A V L V K E - - - - - S G P L E S E A M V M			
Lactococ	R I I D A M K K A E Y M S E - - - - - R I G E V Y T G T I T			
	970	980	990	
SC-Dis3	K V F N N G I V V L V P K F G Y E G L I R L D N L T E - - D P			
Spom1.prot	R I S G N S I D V Y I S E Y G I S N R V D L S S D D R - I K S			
CE	G V M D L S F D V L I V E Y G V V K R V Y V D K M K R - - - D			
HS	G I L K Q A F D V L V L R Y G V Q K R I Y C N A L A L R S H H			
MM	G V L N Q A F D V L V L R F G V Q K R I Y C N A L A L R S Y S			
Lactococ	G L Q K F G I F V A L D N T - V E G L I R V P N I H T - G T T			
	1000	1010	1020	
SC-Dis3	N S A A F D E V E Y K L T F V P T N S D - - - K P - - R D V Y			
Spom1.prot	F I V A P D D S S V K I T L F D D S Q K - - - - - T T A			
CE	F N K S T E K L T I Y W P A D P N A E S G N R E E F S S S I Q			
HS	F Q K V G K K P E L T L V W E P E D M E - - - Q E P A Q Q V I T			
MM	F Q K V G K K P E L T L V W E P D D L E - - - E E P T Q Q V I T			
Lactococ	E E L E F D E E A S I F T G K K S E T V - - - - - Y Q I G			
	1030	1040	1050	
SC-Dis3	V F D K V E V Q V R S V M D P I T S K R K A E L L - - - - -			
Spom1.prot	L T D R F Q V Y L Y S D Y S R T F F S I R C S L V - - - - -			
CE	M C N V V Y V I L V P Y K S I E V S A T I V R P S L E Q - - R			
HS	I F S L V E Y V L Q A E S T A L K Y S A I L K R P G T Q G H L			
MM	I F S L V D V V L Q A E A T A L K Y S A I L K R P - - - - -			
Lactococ	Q E I K I R V I A A N K R K G T V D F E Q I A P E - - - - -			
	1060	1070		
SC-Dis3	- - - - L K - - - - -			
Spom1.prot	- - - - S L N - - - - -			
CE	N I L K S T L K D M K E T G S T I L Q			
HS	G P E K E E E E S D G E P E D S S T S			
MM	G L E K A S D E - - - E P E D - - - -			
Lactococ	- - - - - - - - - - -			



Figur 3



Figur 4

